(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 May 2001 (17.05.2001)

PCT

(10) International Publication Number
WO 01/34802 A2

(51) International Patent Classification[7]: C12N 15/12, 15/62, 15/11, 1/21, 5/10, C07K 14/47, 16/18, 19/00, A61K 38/17, 31/70, 39/395, 48/00, G01N 33/68, C12Q 1/68

(21) International Application Number: PCT/US00/30904

(22) International Filing Date:
9 November 2000 (09.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/439,313	12 November 1999 (12.11.1999)	US
09/443,686	18 November 1999 (18.11.1999)	US

(71) Applicant *(for all designated States except US)*: **CORIXA CORPORATION** [US/US]; Suite 200, 1124 Columbia Street, Seattle, WA 98104 (US).

(72) Inventors; and

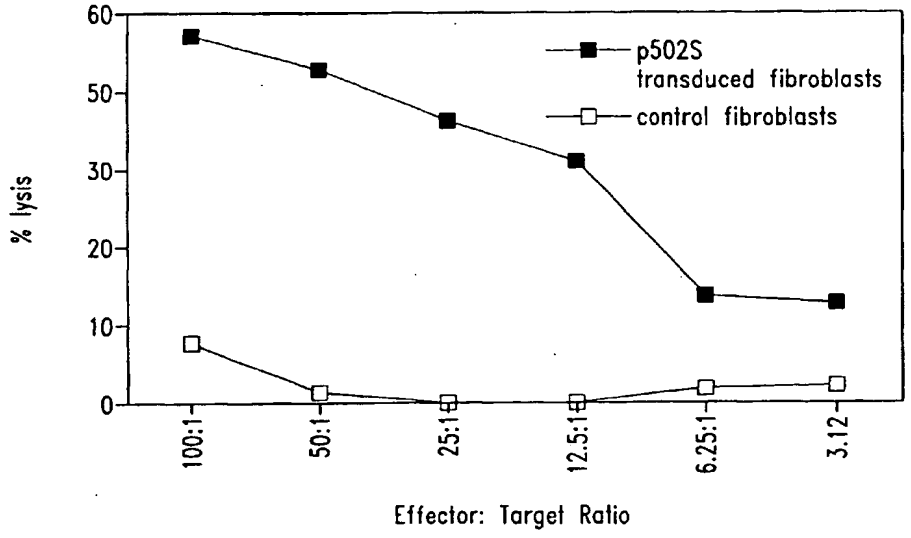
(75) Inventors/Applicants *(for US only)*: **XU, Jiangchun** [US/US]; 15805 SE 43rd Place, Bellevue, WA 98006 (US). **DILLON, Davin, C.** [US/US]; 18112 NW Montreux Drive, Issaquah, WA 98027 (US). **MITCHAM, Jennifer, L.** [US/US]; 16677 NE 88th Street, Redmond, WA 98052 (US). **HARLOCKER, Susan, L.** [US/US]; 7522 - 13th Avenue W., Seattle, WA 98117 (US). **JIANG, Yuqiu** [CN/US]; 5001 South 232nd Street, Kent, WA 98032 (US). **REED, Steven, G.** [US/US]; 2843 - 122nd Place NE, Bellevue, WA 98005 (US). **KALOS, Michael, D.** [US/US]; 8116 Dayton Ave. N., Seattle, WA 98103 (US). **RETTER, Marc, W.** [US/US]; 33402 NE 43rd Place, Carnation, WA 98014 (US). **STOLK, John, A.** [US/US]; 7436 Northeast 144th Place, Bothell, WA 98011 (US). **DAY, Craig, H.** [US/US]; 11501 Stone Ave. N., C122, Seattle, WA 98133-8317 (US). **SKEIKY, Yasir, A.W.** [CA/US]; 15106 SE 47th Place, Bellevue, WA 98006 (US). **WANG, Aijun** [CN/US]; 3106 213th Place SE, Issaquah, WA 98029 (US).

(74) Agents: **POTTER, Jane, E., R.**; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 et al. (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,

[Continued on next page]

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF PROSTATE CANCER



(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, such as prostate cancer, are disclosed. Compositions may comprise one or more prostate-specific proteins, immunogenic portions thereof, or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a prostate-specific protein, or a T cell that is specific for cells expressing such a protein. Such compositions may be used, for example, for the prevention and treatment of diseases such as prostate cancer. Diagnostic methods based on detecting a prostate-specific protein, or mRNA encoding such a protein, in a sample are also provided.

WO 01/34802 A2

DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF PROSTATE CANCER

TECHNICAL FIELD

The present invention relates generally to therapy and diagnosis of cancer, such as prostate cancer. The invention is more specifically related to polypeptides comprising at least a portion of a prostate-specific protein, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides may be used in vaccines and pharmaceutical compositions for prevention and treatment of prostate cancer, and for the diagnosis and monitoring of such cancers.

BACKGROUND OF THE INVENTION

Prostate cancer is the most common form of cancer among males, with an estimated incidence of 30% in men over the age of 50. Overwhelming clinical evidence shows that human prostate cancer has the propensity to metastasize to bone, and the disease appears to progress inevitably from androgen dependent to androgen refractory status, leading to increased patient mortality. This prevalent disease is currently the second leading cause of cancer death among men in the U.S.

In spite of considerable research into therapies for the disease, prostate cancer remains difficult to treat. Commonly, treatment is based on surgery and/or radiation therapy, but these methods are ineffective in a significant percentage of cases. Two previously identified prostate specific proteins - prostate specific antigen (PSA) and prostatic acid phosphatase (PAP) - have limited therapeutic and diagnostic potential. For example, PSA levels do not always correlate well with the presence of prostate cancer, being positive in a percentage of non-prostate cancer cases, including benign prostatic hyperplasia (BPH). Furthermore, PSA measurements correlate with prostate volume, and do not indicate the level of metastasis.

In spite of considerable research into therapies for these and other cancers, prostate cancer remains difficult to diagnose and treat effectively. Accordingly, there is a need in the art for improved methods for detecting and treating such cancers. The present invention fulfills these needs and further provides other related advantages.

SUMMARY OF THE INVENTION

Briefly stated, the present invention provides compositions and methods for the

diagnosis and therapy of cancer, such as prostate cancer. In one aspect, the present invention provides polypeptides comprising at least a portion of a prostate-specific protein, or a variant thereof. Certain portions and other variants are immunogenic, such that the ability of the variant to react with antigen-specific antisera is not substantially diminished. Within certain embodiments, the polypeptide comprises at least an immunogenic portion of a prostate-specific protein, or a variant thereof, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (a) sequences recited in any one of SEQ ID NOs:1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382,384-476, 524, 526, 530, 531, 533, 535 and 536; (b) sequences that hybridize to any of the foregoing sequences under moderately stringent conditions; and (c) complements of any of the sequence of (a) or (b). In certain specific embodiments, such a polypeptide comprises at least a portion, or variant thereof, of a protein that includes an amino acid sequence selected from the group consisting of sequences recited in any one of SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-550.

The present invention further provides polynucleotides that encode a polypeptide as described above, or a portion thereof (such as a portion encoding at least 15 amino acid residues of a prostate-specific protein), expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, vaccines for prophylactic or therapeutic use are provided. Such vaccines comprise a polypeptide or polynucleotide as described above and an immunostimulant.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a prostate-specific protein; and (b) a physiologically acceptable carrier. In certain embodiments, the present invention provides monoclonal antibodies that specifically bind to an amino acid sequence selected from the group consisting of SEQ ID NO: 496, 504, 505, 509-517, 522 and 541-550, together with monoclonal antibodies comprising a complementarity determining region selected from the group consisting of SEQ ID NO: 502, 503 and 506-508.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

Within related aspects, vaccines are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) an immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins.

Within related aspects, pharmaceutical compositions comprising a fusion protein, or a polynucleotide encoding a fusion protein, in combination with a physiologically acceptable carrier are provided.

Vaccines are further provided, within other aspects, that comprise a fusion protein, or a polynucleotide encoding a fusion protein, in combination with an immunostimulant.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a pharmaceutical composition or vaccine as recited above.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a prostate-specific protein, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a prostate-specific protein, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating $CD4^+$ and/or $CD8^+$ T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of a prostate-specific protein; (ii) a polynucleotide encoding such a polypeptide; and (iii) an antigen-presenting cell that expressed such a polypeptide; and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for determining the presence or absence of a cancer in a patient, comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody. The cancer may be prostate cancer.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of: (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a prostate-specific protein; (b) detecting in the sample a level of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain

embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

In related aspects, methods are provided for monitoring the progression of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a prostate-specific protein; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description and attached drawings. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

BRIEF DESCRIPTION OF THE DRAWINGS AND SEQUENCE IDENTIFIERS

Figure 1 illustrates the ability of T cells to kill fibroblasts expressing the representative prostate-specific polypeptide P502S, as compared to control fibroblasts. The percentage lysis is shown as a series of effector:target ratios, as indicated.

Figures 2A and 2B illustrate the ability of T cells to recognize cells expressing the representative prostate-specific polypeptide P502S. In each case, the number of γ-interferon spots is shown for different numbers of responders. In Figure 2A, data is presented for fibroblasts pulsed with the P2S-12 peptide, as compared to fibroblasts pulsed with a control E75 peptide. In Figure 2B, data is presented for fibroblasts expressing P502S, as compared to fibroblasts expressing HER-2/*neu*.

Figure 3 represents a peptide competition binding assay showing that the P1S#10 peptide, derived from P501S, binds HLA-A2. Peptide P1S#10 inhibits HLA-A2 restricted presentation of fluM58 peptide to CTL clone D150M58 in TNF release bioassay. D150M58 CTL is specific for the HLA-A2 binding influenza matrix peptide fluM58.

Figure 4 illustrates the ability of T cell lines generated from P1S#10 immunized mice to specifically lyse P1S#10-pulsed Jurkat A2Kb targets and P501S-transduced Jurkat A2Kb targets, as compared to EGFP-transduced Jurkat A2Kb. The percent lysis is shown as a series of effector to target ratios, as indicated.

Figure 5 illustrates the ability of a T cell clone to recognize and specifically lyse Jurkat A2Kb cells expressing the representative prostate-specific polypeptide P501S, thereby demonstrating that the P1S#10 peptide may be a naturally processed epitope of the P501S polypeptide.

Figures 6A and 6B are graphs illustrating the specificity of a $CD8^+$ cell line (3A-1) for a representative prostate-specific antigen (P501S). Figure 6A shows the results of a ^{51}Cr release assay. The percent specific lysis is shown as a series of effector:target ratios, as indicated. Figure 6B shows the production of interferon-gamma by 3A-1 cells stimulated with autologous B-LCL transduced with P501S, at varying effector:target rations as indicated.

Figure 7 is a Western blot showing the expression of P501S in baculovirus.

Figure 8 illustrates the results of epitope mapping studies on P501S.

Figure 9 is a schematic representation of the P501S protein showing the location of transmembrane domains and predicted intracellular and extracellular domains.

Figure 10 is a genomic map showing the location of the prostate genes P775P, P704P, B305D, P712P and P774P within the Cat Eye Syndrome region of chromosome 22q11.2

Figure 11 shows the results of an ELISA assay of antibody specificity to P501S peptides.

SEQ ID NO: 1 is the determined cDNA sequence for F1-13

SEQ ID NO: 2 is the determined 3' cDNA sequence for F1-12

SEQ ID NO: 3 is the determined 5' cDNA sequence for F1-12

SEQ ID NO: 4 is the determined 3' cDNA sequence for F1-16

SEQ ID NO: 5 is the determined 3' cDNA sequence for H1-1

SEQ ID NO: 6 is the determined 3' cDNA sequence for H1-9

SEQ ID NO: 7 is the determined 3' cDNA sequence for H1-4

SEQ ID NO: 8 is the determined 3' cDNA sequence for J1-17

SEQ ID NO: 9 is the determined 5' cDNA sequence for J1-17

SEQ ID NO: 10 is the determined 3' cDNA sequence for L1-12

SEQ ID NO: 11 is the determined 5' cDNA sequence for L1-12

SEQ ID NO: 12 is the determined 3' cDNA sequence for N1-1862

SEQ ID NO: 13 is the determined 5' cDNA sequence for N1-1862

SEQ ID NO: 14 is the determined 3' cDNA sequence for J1-13

SEQ ID NO: 15 is the determined 5' cDNA sequence for J1-13

SEQ ID NO: 16 is the determined 3' cDNA sequence for J1-19

SEQ ID NO: 17 is the determined 5' cDNA sequence for J1-19

SEQ ID NO: 18 is the determined 3' cDNA sequence for J1-25

SEQ ID NO: 19 is the determined 5' cDNA sequence for J1-25

SEQ ID NO: 20 is the determined 5' cDNA sequence for J1-24

SEQ ID NO: 21 is the determined 3' cDNA sequence for J1-24

SEQ ID NO: 22 is the determined 5' cDNA sequence for K1-58

SEQ ID NO: 23 is the determined 3' cDNA sequence for K1-58

SEQ ID NO: 24 is the determined 5' cDNA sequence for K1-63

SEQ ID NO: 25 is the determined 3' cDNA sequence for K1-63

SEQ ID NO: 26 is the determined 5' cDNA sequence for L1-4

SEQ ID NO: 27 is the determined 3' cDNA sequence for L1-4

SEQ ID NO: 28 is the determined 5' cDNA sequence for L1-14

SEQ ID NO: 29 is the determined 3' cDNA sequence for L1-14

SEQ ID NO: 30 is the determined 3' cDNA sequence for J1-12

SEQ ID NO: 31 is the determined 3' cDNA sequence for J1-16

SEQ ID NO: 32 is the determined 3' cDNA sequence for J1-21

SEQ ID NO: 33 is the determined 3' cDNA sequence for K1-48

SEQ ID NO: 34 is the determined 3' cDNA sequence for K1-55

SEQ ID NO: 35 is the determined 3' cDNA sequence for L1-2

SEQ ID NO: 36 is the determined 3' cDNA sequence for L1-6

SEQ ID NO: 37 is the determined 3' cDNA sequence for N1-1858

SEQ ID NO: 38 is the determined 3' cDNA sequence for N1-1860

SEQ ID NO: 39 is the determined 3' cDNA sequence for N1-1861

SEQ ID NO: 40 is the determined 3' cDNA sequence for N1-1864

SEQ ID NO: 41 is the determined cDNA sequence for P5

SEQ ID NO: 42 is the determined cDNA sequence for P8

SEQ ID NO: 43 is the determined cDNA sequence for P9

SEQ ID NO: 44 is the determined cDNA sequence for P18

SEQ ID NO: 45 is the determined cDNA sequence for P20

SEQ ID NO: 46 is the determined cDNA sequence for P29

SEQ ID NO: 47 is the determined cDNA sequence for P30

SEQ ID NO: 48 is the determined cDNA sequence for P34

SEQ ID NO: 49 is the determined cDNA sequence for P36

SEQ ID NO: 50 is the determined cDNA sequence for P38

SEQ ID NO: 51 is the determined cDNA sequence for P39

SEQ ID NO: 52 is the determined cDNA sequence for P42

SEQ ID NO: 53 is the determined cDNA sequence for P47

SEQ ID NO: 54 is the determined cDNA sequence for P49

SEQ ID NO: 55 is the determined cDNA sequence for P50

SEQ ID NO: 56 is the determined cDNA sequence for P53

SEQ ID NO: 57 is the determined cDNA sequence for P55

SEQ ID NO: 58 is the determined cDNA sequence for P60

SEQ ID NO: 59 is the determined cDNA sequence for P64

SEQ ID NO: 60 is the determined cDNA sequence for P65

SEQ ID NO: 61 is the determined cDNA sequence for P73

SEQ ID NO: 62 is the determined cDNA sequence for P75

SEQ ID NO: 63 is the determined cDNA sequence for P76

SEQ ID NO: 64 is the determined cDNA sequence for P79

SEQ ID NO: 65 is the determined cDNA sequence for P84

SEQ ID NO: 66 is the determined cDNA sequence for P68

SEQ ID NO: 67 is the determined cDNA sequence for P80

SEQ ID NO: 68 is the determined cDNA sequence for P82

SEQ ID NO: 69 is the determined cDNA sequence for U1-3064

SEQ ID NO: 70 is the determined cDNA sequence for U1-3065

SEQ ID NO: 71 is the determined cDNA sequence for V1-3692

SEQ ID NO: 72 is the determined cDNA sequence for 1A-3905

SEQ ID NO: 73 is the determined cDNA sequence for V1-3686

SEQ ID NO: 74 is the determined cDNA sequence for R1-2330

SEQ ID NO: 75 is the determined cDNA sequence for 1B-3976

SEQ ID NO: 76 is the determined cDNA sequence for V1-3679

SEQ ID NO: 77 is the determined cDNA sequence for1G-4736

SEQ ID NO: 78 is the determined cDNA sequence for 1G-4738

SEQ ID NO: 79 is the determined cDNA sequence for 1G-4741

SEQ ID NO: 80 is the determined cDNA sequence for 1G-4744

SEQ ID NO: 81 is the determined cDNA sequence for 1G-4734

SEQ ID NO: 82 is the determined cDNA sequence for 1H-4774

SEQ ID NO: 83 is the determined cDNA sequence for 1H-4781

SEQ ID NO: 84 is the determined cDNA sequence for 1H-4785

SEQ ID NO: 85 is the determined cDNA sequence for 1H-4787

SEQ ID NO: 86 is the determined cDNA sequence for 1H-4796

SEQ ID NO: 87 is the determined cDNA sequence for 1I-4807

SEQ ID NO: 88 is the determined cDNA sequence for 1I-4810

SEQ ID NO: 89 is the determined cDNA sequence for 1I-4811

SEQ ID NO: 90 is the determined cDNA sequence for 1J-4876

SEQ ID NO: 91 is the determined cDNA sequence for 1K-4884

SEQ ID NO: 92 is the determined cDNA sequence for 1K-4896

SEQ ID NO: 93 is the determined cDNA sequence for 1G-4761

SEQ ID NO: 94 is the determined cDNA sequence for 1G-4762

SEQ ID NO: 95 is the determined cDNA sequence for 1H-4766

SEQ ID NO: 96 is the determined cDNA sequence for 1H-4770

SEQ ID NO: 97 is the determined cDNA sequence for 1H-4771

SEQ ID NO: 98 is the determined cDNA sequence for 1H-4772

SEQ ID NO: 99 is the determined cDNA sequence for 1D-4297

SEQ ID NO: 100 is the determined cDNA sequence for 1D-4309

SEQ ID NO: 101 is the determined cDNA sequence for 1D.1-4278

SEQ ID NO: 102 is the determined cDNA sequence for 1D-4288

SEQ ID NO: 103 is the determined cDNA sequence for 1D-4283

SEQ ID NO: 104 is the determined cDNA sequence for 1D-4304

SEQ ID NO: 105 is the determined cDNA sequence for 1D-4296

SEQ ID NO: 106 is the determined cDNA sequence for 1D-4280

SEQ ID NO: 107 is the determined full length cDNA sequence for F1-12 (also referred to as P504S)

SEQ ID NO: 108 is the predicted amino acid sequence for F1-12

SEQ ID NO: 109 is the determined full length cDNA sequence for J1-17

SEQ ID NO: 110 is the determined full length cDNA sequence for L1-12 (also referred to as P501S)

SEQ ID NO: 111 is the determined full length cDNA sequence for N1-1862 (also referred to as P503S)

SEQ ID NO: 112 is the predicted amino acid sequence for J1-17

SEQ ID NO: 113 is the predicted amino acid sequence for L1-12 (also referred to as P501S)

SEQ ID NO: 114 is the predicted amino acid sequence for N1-1862 (also referred to as P503S)

SEQ ID NO: 115 is the determined cDNA sequence for P89

SEQ ID NO: 116 is the determined cDNA sequence for P90

SEQ ID NO: 117 is the determined cDNA sequence for P92

SEQ ID NO: 118 is the determined cDNA sequence for P95

SEQ ID NO: 119 is the determined cDNA sequence for P98

SEQ ID NO: 120 is the determined cDNA sequence for P102

SEQ ID NO: 121 is the determined cDNA sequence for P110

SEQ ID NO: 122 is the determined cDNA sequence for P111

SEQ ID NO: 123 is the determined cDNA sequence for P114

SEQ ID NO: 124 is the determined cDNA sequence for P115

SEQ ID NO: 125 is the determined cDNA sequence for P116

SEQ ID NO: 126 is the determined cDNA sequence for P124

SEQ ID NO: 127 is the determined cDNA sequence for P126

SEQ ID NO: 128 is the determined cDNA sequence for P130

SEQ ID NO: 129 is the determined cDNA sequence for P133

SEQ ID NO: 130 is the determined cDNA sequence for P138

SEQ ID NO: 131 is the determined cDNA sequence for P143

SEQ ID NO: 132 is the determined cDNA sequence for P151

SEQ ID NO: 133 is the determined cDNA sequence for P156

SEQ ID NO: 134 is the determined cDNA sequence for P157

SEQ ID NO: 135 is the determined cDNA sequence for P166

SEQ ID NO: 136 is the determined cDNA sequence for P176

SEQ ID NO: 137 is the determined cDNA sequence for P178

SEQ ID NO: 138 is the determined cDNA sequence for P179

SEQ ID NO: 139 is the determined cDNA sequence for P185

SEQ ID NO: 140 is the determined cDNA sequence for P192

SEQ ID NO: 141 is the determined cDNA sequence for P201

SEQ ID NO: 142 is the determined cDNA sequence for P204

SEQ ID NO: 143 is the determined cDNA sequence for P208

SEQ ID NO: 144 is the determined cDNA sequence for P211

SEQ ID NO: 145 is the determined cDNA sequence for P213

SEQ ID NO: 146 is the determined cDNA sequence for P219

SEQ ID NO: 147 is the determined cDNA sequence for P237

SEQ ID NO: 148 is the determined cDNA sequence for P239

SEQ ID NO: 149 is the determined cDNA sequence for P248

SEQ ID NO: 150 is the determined cDNA sequence for P251

SEQ ID NO: 151 is the determined cDNA sequence for P255

SEQ ID NO: 152 is the determined cDNA sequence for P256

SEQ ID NO: 153 is the determined cDNA sequence for P259

SEQ ID NO: 154 is the determined cDNA sequence for P260

SEQ ID NO: 155 is the determined cDNA sequence for P263

SEQ ID NO: 156 is the determined cDNA sequence for P264

SEQ ID NO: 157 is the determined cDNA sequence for P266

SEQ ID NO: 158 is the determined cDNA sequence for P270

SEQ ID NO: 159 is the determined cDNA sequence for P272

SEQ ID NO: 160 is the determined cDNA sequence for P278

SEQ ID NO: 161 is the determined cDNA sequence for P105

SEQ ID NO: 162 is the determined cDNA sequence for P107

SEQ ID NO: 163 is the determined cDNA sequence for P137

SEQ ID NO: 164 is the determined cDNA sequence for P194

SEQ ID NO: 165 is the determined cDNA sequence for P195

SEQ ID NO: 166 is the determined cDNA sequence for P196

SEQ ID NO: 167 is the determined cDNA sequence for P220

SEQ ID NO: 168 is the determined cDNA sequence for P234

SEQ ID NO: 169 is the determined cDNA sequence for P235

SEQ ID NO: 170 is the determined cDNA sequence for P243

SEQ ID NO: 171 is the determined cDNA sequence for P703P-DE1

SEQ ID NO: 172 is the predicted amino acid sequence for P703P-DE1

SEQ ID NO: 173 is the determined cDNA sequence for P703P-DE2

SEQ ID NO: 174 is the determined cDNA sequence for P703P-DE6

SEQ ID NO: 175 is the determined cDNA sequence for P703P-DE13

SEQ ID NO: 176 is the predicted amino acid sequence for P703P-DE13

SEQ ID NO: 177 is the determined cDNA sequence for P703P-DE14

SEQ ID NO: 178 is the predicted amino acid sequence for P703P-DE14

SEQ ID NO: 179 is the determined extended cDNA sequence for 1G-4736

SEQ ID NO: 180 is the determined extended cDNA sequence for 1G-4738

SEQ ID NO: 181 is the determined extended cDNA sequence for 1G-4741

SEQ ID NO: 182 is the determined extended cDNA sequence for 1G-4744

SEQ ID NO: 183 is the determined extended cDNA sequence for 1H-4774

SEQ ID NO: 184 is the determined extended cDNA sequence for 1H-4781

SEQ ID NO: 185 is the determined extended cDNA sequence for 1H-4785

SEQ ID NO: 186 is the determined extended cDNA sequence for 1H-4787

SEQ ID NO: 187 is the determined extended cDNA sequence for 1H-4796

SEQ ID NO: 188 is the determined extended cDNA sequence for 1I-4807

SEQ ID NO: 189 is the determined 3' cDNA sequence for 1I-4810

SEQ ID NO: 190 is the determined 3' cDNA sequence for 1I-4811

SEQ ID NO: 191 is the determined extended cDNA sequence for 1J-4876

SEQ ID NO: 192 is the determined extended cDNA sequence for 1K-4884

SEQ ID NO: 193 is the determined extended cDNA sequence for 1K-4896

SEQ ID NO: 194 is the determined extended cDNA sequence for 1G-4761

SEQ ID NO: 195 is the determined extended cDNA sequence for 1G-4762

SEQ ID NO: 196 is the determined extended cDNA sequence for 1H-4766

SEQ ID NO: 197 is the determined 3' cDNA sequence for 1H-4770

SEQ ID NO: 198 is the determined 3' cDNA sequence for 1H-4771

SEQ ID NO: 199 is the determined extended cDNA sequence for 1H-4772

SEQ ID NO: 200 is the determined extended cDNA sequence for 1D-4309

SEQ ID NO: 201 is the determined extended cDNA sequence for 1D.1-4278

SEQ ID NO: 202 is the determined extended cDNA sequence for 1D-4288

SEQ ID NO: 203 is the determined extended cDNA sequence for 1D-4283

SEQ ID NO: 204 is the determined extended cDNA sequence for 1D-4304

SEQ ID NO: 205 is the determined extended cDNA sequence for 1D-4296

SEQ ID NO: 206 is the determined extended cDNA sequence for 1D-4280

SEQ ID NO: 207 is the determined cDNA sequence for 10-d8fwd

SEQ ID NO: 208 is the determined cDNA sequence for 10-H10con

SEQ ID NO: 209 is the determined cDNA sequence for 11-C8rev

SEQ ID NO: 210 is the determined cDNA sequence for 7.g6fwd

SEQ ID NO: 211 is the determined cDNA sequence for 7.g6rev

SEQ ID NO: 212 is the determined cDNA sequence for 8-b5fwd

SEQ ID NO: 213 is the determined cDNA sequence for 8-b5rev

SEQ ID NO: 214 is the determined cDNA sequence for 8-b6fwd

SEQ ID NO: 215 is the determined cDNA sequence for 8-b6 rev

SEQ ID NO: 216 is the determined cDNA sequence for 8-d4fwd

SEQ ID NO: 217 is the determined cDNA sequence for 8-d9rev

SEQ ID NO: 218 is the determined cDNA sequence for 8-g3fwd

SEQ ID NO: 219 is the determined cDNA sequence for 8-g3rev

SEQ ID NO: 220 is the determined cDNA sequence for 8-h11rev

SEQ ID NO: 221 is the determined cDNA sequence for g-f12fwd

SEQ ID NO: 222 is the determined cDNA sequence for g-f3rev

SEQ ID NO: 223 is the determined cDNA sequence for P509S

SEQ ID NO: 224 is the determined cDNA sequence for P510S

SEQ ID NO: 225 is the determined cDNA sequence for P703DE5

SEQ ID NO: 226 is the determined cDNA sequence for 9-A11

SEQ ID NO: 227 is the determined cDNA sequence for 8-C6

SEQ ID NO: 228 is the determined cDNA sequence for 8-H7

SEQ ID NO: 229 is the determined cDNA sequence for JPTPN13

SEQ ID NO: 230 is the determined cDNA sequence for JPTPN14

SEQ ID NO: 231 is the determined cDNA sequence for JPTPN23

SEQ ID NO: 232 is the determined cDNA sequence for JPTPN24

SEQ ID NO: 233 is the determined cDNA sequence for JPTPN25

SEQ ID NO: 234 is the determined cDNA sequence for JPTPN30

SEQ ID NO: 235 is the determined cDNA sequence for JPTPN34

SEQ ID NO: 236 is the determined cDNA sequence for PTPN35

SEQ ID NO: 237 is the determined cDNA sequence for JPTPN36

SEQ ID NO: 238 is the determined cDNA sequence for JPTPN38

SEQ ID NO: 239 is the determined cDNA sequence for JPTPN39

SEQ ID NO: 240 is the determined cDNA sequence for JPTPN40

SEQ ID NO: 241 is the determined cDNA sequence for JPTPN41

SEQ ID NO: 242 is the determined cDNA sequence for JPTPN42

SEQ ID NO: 243 is the determined cDNA sequence for JPTPN45

SEQ ID NO: 244 is the determined cDNA sequence for JPTPN46

SEQ ID NO: 245 is the determined cDNA sequence for JPTPN51

SEQ ID NO: 246 is the determined cDNA sequence for JPTPN56

SEQ ID NO: 247 is the determined cDNA sequence for PTPN64

SEQ ID NO: 248 is the determined cDNA sequence for JPTPN65

SEQ ID NO: 249 is the determined cDNA sequence for JPTPN67

SEQ ID NO: 250 is the determined cDNA sequence for JPTPN76

SEQ ID NO: 251 is the determined cDNA sequence for JPTPN84

SEQ ID NO: 252 is the determined cDNA sequence for JPTPN85

SEQ ID NO: 253 is the determined cDNA sequence for JPTPN86

SEQ ID NO: 254 is the determined cDNA sequence for JPTPN87

SEQ ID NO: 255 is the determined cDNA sequence for JPTPN88

SEQ ID NO: 256 is the determined cDNA sequence for JP1F1

SEQ ID NO: 257 is the determined cDNA sequence for JP1F2

SEQ ID NO: 258 is the determined cDNA sequence for JP1C2

SEQ ID NO: 259 is the determined cDNA sequence for JP1B1

SEQ ID NO: 260 is the determined cDNA sequence for JP1B2

SEQ ID NO: 261 is the determined cDNA sequence for JP1D3

SEQ ID NO: 262 is the determined cDNA sequence for JP1A4

SEQ ID NO: 263 is the determined cDNA sequence for JP1F5

SEQ ID NO: 264 is the determined cDNA sequence for JP1E6

SEQ ID NO: 265 is the determined cDNA sequence for JP1D6

SEQ ID NO: 266 is the determined cDNA sequence for JP1B5

SEQ ID NO: 267 is the determined cDNA sequence for JP1A6

SEQ ID NO: 268 is the determined cDNA sequence for JP1E8

SEQ ID NO: 269 is the determined cDNA sequence for JP1D7

SEQ ID NO: 270 is the determined cDNA sequence for JP1D9

SEQ ID NO: 271 is the determined cDNA sequence for JP1C10

SEQ ID NO: 272 is the determined cDNA sequence for JP1A9

SEQ ID NO: 273 is the determined cDNA sequence for JP1F12

SEQ ID NO: 274 is the determined cDNA sequence for JP1E12

SEQ ID NO: 275 is the determined cDNA sequence for JP1D11

SEQ ID NO: 276 is the determined cDNA sequence for JP1C11

SEQ ID NO: 277 is the determined cDNA sequence for JP1C12

SEQ ID NO: 278 is the determined cDNA sequence for JP1B12

SEQ ID NO: 279 is the determined cDNA sequence for JP1A12

SEQ ID NO: 280 is the determined cDNA sequence for JP8G2

SEQ ID NO: 281 is the determined cDNA sequence for JP8H1

SEQ ID NO: 282 is the determined cDNA sequence for JP8H2

SEQ ID NO: 283 is the determined cDNA sequence for JP8A3

SEQ ID NO: 284 is the determined cDNA sequence for JP8A4

SEQ ID NO: 285 is the determined cDNA sequence for JP8C3

SEQ ID NO: 286 is the determined cDNA sequence for JP8G4

SEQ ID NO: 287 is the determined cDNA sequence for JP8B6

SEQ ID NO: 288 is the determined cDNA sequence for JP8D6

SEQ ID NO: 289 is the determined cDNA sequence for JP8F5

SEQ ID NO: 290 is the determined cDNA sequence for JP8A8

SEQ ID NO: 291 is the determined cDNA sequence for JP8C7

SEQ ID NO: 292 is the determined cDNA sequence for JP8D7

SEQ ID NO: 293 is the determined cDNA sequence for P8D8

SEQ ID NO: 294 is the determined cDNA sequence for JP8E7

SEQ ID NO: 295 is the determined cDNA sequence for JP8F8

SEQ ID NO: 296 is the determined cDNA sequence for JP8G8

SEQ ID NO: 297 is the determined cDNA sequence for JP8B10

SEQ ID NO: 298 is the determined cDNA sequence for JP8C10

SEQ ID NO: 299 is the determined cDNA sequence for JP8E9

SEQ ID NO: 300 is the determined cDNA sequence for JP8E10

SEQ ID NO: 301 is the determined cDNA sequence for JP8F9

SEQ ID NO: 302 is the determined cDNA sequence for JP8H9

SEQ ID NO: 303 is the determined cDNA sequence for JP8C12

SEQ ID NO: 304 is the determined cDNA sequence for JP8E11

SEQ ID NO: 305 is the determined cDNA sequence for JP8E12

SEQ ID NO: 306 is the amino acid sequence for the peptide PS2#12

SEQ ID NO: 307 is the determined cDNA sequence for P711P

SEQ ID NO: 308 is the determined cDNA sequence for P712P

SEQ ID NO: 309 is the determined cDNA sequence for CLONE23

SEQ ID NO: 310 is the determined cDNA sequence for P774P

SEQ ID NO: 311 is the determined cDNA sequence for P775P

SEQ ID NO: 312 is the determined cDNA sequence for P715P

SEQ ID NO: 313 is the determined cDNA sequence for P710P

SEQ ID NO: 314 is the determined cDNA sequence for P767P

SEQ ID NO: 315 is the determined cDNA sequence for P768P

SEQ ID NO: 316-325 are the determined cDNA sequences of previously isolated genes

SEQ ID NO: 326 is the determined cDNA sequence for P703PDE5

SEQ ID NO: 327 is the predicted amino acid sequence for P703PDE5

SEQ ID NO: 328 is the determined cDNA sequence for P703P6.26

SEQ ID NO: 329 is the predicted amino acid sequence for P703P6.26

SEQ ID NO: 330 is the determined cDNA sequence for P703PX-23

SEQ ID NO: 331 is the predicted amino acid sequence for P703PX-23

SEQ ID NO: 332 is the determined full length cDNA sequence for P509S

SEQ ID NO: 333 is the determined extended cDNA sequence for P707P (also referred to as 11-C9)

SEQ ID NO: 334 is the determined cDNA sequence for P714P

SEQ ID NO: 335 is the determined cDNA sequence for P705P (also referred to as 9-F3)

SEQ ID NO: 336 is the predicted amino acid sequence for P705P

SEQ ID NO: 337 is the amino acid sequence of the peptide P1S#10

SEQ ID NO: 338 is the amino acid sequence of the peptide p5

SEQ ID NO: 339 is the predicted amino acid sequence of P509S

SEQ ID NO: 340 is the determined cDNA sequence for P778P

SEQ ID NO: 341 is the determined cDNA sequence for P786P

SEQ ID NO: 342 is the determined cDNA sequence for P789P

SEQ ID NO: 343 is the determined cDNA sequence for a clone showing homology to Homo sapiens MM46 mRNA

SEQ ID NO: 344 is the determined cDNA sequence for a clone showing homology to Homo sapiens TNF-alpha stimulated ABC protein (ABC50) mRNA

SEQ ID NO: 345 is the determined cDNA sequence for a clone showing homology to Homo sapiens mRNA for E-cadherin

SEQ ID NO: 346 is the determined cDNA sequence for a clone showing homology to Human nuclear-encoded mitochondrial serine hydroxymethyltransferase (SHMT)

SEQ ID NO: 347 is the determined cDNA sequence for a clone showing homology to Homo sapiens natural resistance-associated macrophage protein2 (NRAMP2)

SEQ ID NO: 348 is the determined cDNA sequence for a clone showing homology to Homo sapiens phosphoglucomutase-related protein (PGMRP)

SEQ ID NO: 349 is the determined cDNA sequence for a clone showing homology to Human mRNA for proteosome subunit p40

SEQ ID NO: 350 is the determined cDNA sequence for P777P

SEQ ID NO: 351 is the determined cDNA sequence for P779P

SEQ ID NO: 352 is the determined cDNA sequence for P790P

SEQ ID NO: 353 is the determined cDNA sequence for P784P

SEQ ID NO: 354 is the determined cDNA sequence for P776P

SEQ ID NO: 355 is the determined cDNA sequence for P780P

SEQ ID NO: 356 is the determined cDNA sequence for P544S

SEQ ID NO: 357 is the determined cDNA sequence for P745S

SEQ ID NO: 358 is the determined cDNA sequence for P782P

SEQ ID NO: 359 is the determined cDNA sequence for P783P

SEQ ID NO: 360 is the determined cDNA sequence for unknown 17984

SEQ ID NO: 361 is the determined cDNA sequence for P787P

SEQ ID NO: 362 is the determined cDNA sequence for P788P

SEQ ID NO: 363 is the determined cDNA sequence for unknown 17994

SEQ ID NO: 364 is the determined cDNA sequence for P781P

SEQ ID NO: 365 is the determined cDNA sequence for P785P

SEQ ID NO: 366-375 are the determined cDNA sequences for splice variants of B305D.

SEQ ID NO: 376 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 366.

SEQ ID NO: 377 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 372.

SEQ ID NO: 378 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 373.

SEQ ID NO: 379 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 374.

SEQ ID NO: 380 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 375.

SEQ ID NO: 381 is the determined cDNA sequence for B716P.

SEQ ID NO: 382 is the determined full-length cDNA sequence for P711P.

SEQ ID NO: 383 is the predicted amino acid sequence for P711P.

SEQ ID NO: 384 is the cDNA sequence for P1000C.

SEQ ID NO: 385 is the cDNA sequence for CGI-82.

SEQ ID NO:386 is the cDNA sequence for 23320.

SEQ ID NO:387 is the cDNA sequence for CGI-69.

SEQ ID NO:388 is the cDNA sequence for L-iditol-2-dehydrogenase.

SEQ ID NO:389 is the cDNA sequence for 23379.

SEQ ID NO:390 is the cDNA sequence for 23381.

SEQ ID NO:391 is the cDNA sequence for KIAA0122.

SEQ ID NO:392 is the cDNA sequence for 23399.

SEQ ID NO:393 is the cDNA sequence for a previously identified gene.

SEQ ID NO:394 is the cDNA sequence for HCLBP.

SEQ ID NO:395 is the cDNA sequence for transglutaminase.

SEQ ID NO:396 is the cDNA sequence for a previously identified gene.

SEQ ID NO:397 is the cDNA sequence for PAP.

SEQ ID NO:398 is the cDNA sequence for Ets transcription factor PDEF.

SEQ ID NO:399 is the cDNA sequence for hTGR.

SEQ ID NO:400 is the cDNA sequence for KIAA0295.

SEQ ID NO:401 is the cDNA sequence for 22545.

SEQ ID NO:402 is the cDNA sequence for 22547.

SEQ ID NO:403 is the cDNA sequence for 22548.

SEQ ID NO:404 is the cDNA sequence for 22550.

SEQ ID NO:405 is the cDNA sequence for 22551.

SEQ ID NO:406 is the cDNA sequence for 22552.

SEQ ID NO:407 is the cDNA sequence for 22553.

SEQ ID NO:408 is the cDNA sequence for 22558.

SEQ ID NO:409 is the cDNA sequence for 22562.

SEQ ID NO:410 is the cDNA sequence for 22565.

SEQ ID NO:411 is the cDNA sequence for 22567.

SEQ ID NO:412 is the cDNA sequence for 22568.

SEQ ID NO:413 is the cDNA sequence for 22570.

SEQ ID NO:414 is the cDNA sequence for 22571.

SEQ ID NO:415 is the cDNA sequence for 22572.

SEQ ID NO:416 is the cDNA sequence for 22573.

SEQ ID NO:417 is the cDNA sequence for 22573.

SEQ ID NO:418 is the cDNA sequence for 22575.

SEQ ID NO:419 is the cDNA sequence for 22580.

SEQ ID NO:420 is the cDNA sequence for 22581.

SEQ ID NO:421 is the cDNA sequence for 22582.

SEQ ID NO:422 is the cDNA sequence for 22583.

SEQ ID NO:423 is the cDNA sequence for 22584.

SEQ ID NO:424 is the cDNA sequence for 22585.

SEQ ID NO:425 is the cDNA sequence for 22586.

SEQ ID NO:426 is the cDNA sequence for 22587.

SEQ ID NO:427 is the cDNA sequence for 22588.

SEQ ID NO:428 is the cDNA sequence for 22589.

SEQ ID NO:429 is the cDNA sequence for 22590.

SEQ ID NO:430 is the cDNA sequence for 22591.

SEQ ID NO:431 is the cDNA sequence for 22592.

SEQ ID NO:432 is the cDNA sequence for 22593.

SEQ ID NO:433 is the cDNA sequence for 22594.

SEQ ID NO:434 is the cDNA sequence for 22595.

SEQ ID NO:435 is the cDNA sequence for 22596.

SEQ ID NO:436 is the cDNA sequence for 22847.

SEQ ID NO:437 is the cDNA sequence for 22848.

SEQ ID NO:438 is the cDNA sequence for 22849.

SEQ ID NO:439 is the cDNA sequence for 22851.

SEQ ID NO:440 is the cDNA sequence for 22852.

SEQ ID NO:441 is the cDNA sequence for 22853.

SEQ ID NO:442 is the cDNA sequence for 22854.

SEQ ID NO:443 is the cDNA sequence for 22855.

SEQ ID NO:444 is the cDNA sequence for 22856.

SEQ ID NO:445 is the cDNA sequence for 22857.

SEQ ID NO:446 is the cDNA sequence for 23601.

SEQ ID NO:447 is the cDNA sequence for 23602.

SEQ ID NO:448 is the cDNA sequence for 23605.

SEQ ID NO:449 is the cDNA sequence for 23606.

SEQ ID NO:450 is the cDNA sequence for 23612.

SEQ ID NO:451 is the cDNA sequence for 23614.

SEQ ID NO:452 is the cDNA sequence for 23618.

SEQ ID NO:453 is the cDNA sequence for 23622.

SEQ ID NO:454 is the cDNA sequence for folate hydrolase.

SEQ ID NO:455 is the cDNA sequence for LIM protein.

SEQ ID NO:456 is the cDNA sequence for a known gene.

SEQ ID NO:457 is the cDNA sequence for a known gene.

SEQ ID NO:458 is the cDNA sequence for a previously identified gene.

SEQ ID NO:459 is the cDNA sequence for 23045.

SEQ ID NO:460 is the cDNA sequence for 23032.

SEQ ID NO:461 is the cDNA sequence for 23054.

SEQ ID NO:462-467 are cDNA sequences for known genes.

SEQ ID NO:468-471 are cDNA sequences for P710P.

SEQ ID NO:472 is a cDNA sequence for P1001C.

SEQ ID NO: 473 is the determined cDNA sequence for a first splice variant of P775P (referred to as 27505).

SEQ ID NO: 474 is the determined cDNA sequence for a second splice variant of P775P (referred to as 19947).

SEQ ID NO: 475 is the determined cDNA sequence for a third splice variant of P775P (referred to as 19941).

SEQ ID NO: 476 is the determined cDNA sequence for a fourth splice variant of P775P (referred to as 19937).

SEQ ID NO: 477 is a first predicted amino acid sequence encoded by the sequence of SEQ ID NO: 474.

SEQ ID NO: 478 is a second predicted amino acid sequence encoded by the sequence of SEQ ID NO: 474.

SEQ ID NO: 479 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 475.

SEQ ID NO: 480 is a first predicted amino acid sequence encoded by the sequence of SEQ ID NO: 473.

SEQ ID NO: 481 is a second predicted amino acid sequence encoded by the sequence of SEQ ID NO: 473.

SEQ ID NO: 482 is a third predicted amino acid sequence encoded by the sequence of SEQ ID NO: 473.

SEQ ID NO: 483 is a fourth predicted amino acid sequence encoded by the sequence of SEQ ID NO: 473.

SEQ ID NO: 484 is the first 30 amino acids of the *M. tuberculosis* antigen Ra12.

SEQ ID NO: 485 is the PCR primer AW025.

SEQ ID NO: 486 is the PCR primer AW003.

SEQ ID NO: 487 is the PCR primer AW027.

SEQ ID NO: 488 is the PCR primer AW026.

SEQ ID NO: 489-501 are peptides employed in epitope mapping studies.

SEQ ID NO: 502 is the determined cDNA sequence of the complementarity determining region for the anti-P503S monoclonal antibody 20D4.

SEQ ID NO: 503 is the determined cDNA sequence of the complementarity determining region for the anti-P503S monoclonal antibody JA1.

SEQ ID NO: 504 & 505 are peptides employed in epitope mapping studies.

SEQ ID NO: 506 is the determined cDNA sequence of the complementarity determining region for the anti-P703P monoclonal antibody 8H2.

SEQ ID NO: 507 is the determined cDNA sequence of the complementarity determining region for the anti-P703P monoclonal antibody 7H8.

SEQ ID NO: 508 is the determined cDNA sequence of the complementarity determining region for the anti-P703P monoclonal antibody 2D4.

SEQ ID NO: 509-522 are peptides employed in epitope mapping studies.

SEQ ID NO: 523 is a mature form of P703P used to raise antibodies against P703P.SEQ ID NO: 524 is the putative full-length cDNA sequence of P703P.

SEQ ID NO: 525 is the predicted amino acid sequence encoded by SEQ ID NO: 524.

SEQ ID NO: 526 is the full-length cDNA sequence for P790P.

SEQ ID NO: 527 is the predicted amino acid sequence for P790P.

SEQ ID NO: 528 & 529 are PCR primers.

SEQ ID NO: 530 is the cDNA sequence of a splice variant of SEQ ID NO: 366.

SEQ ID NO: 531 is the cDNA sequence of the open reading frame of SEQ ID NO: 530.

SEQ ID NO: 532 is the predicted amino acid encoded by the sequence of SEQ ID NO: 531.

SEQ ID NO: 533 is the DNA sequence of a putative ORF of P775P.

SEQ ID NO: 534 is the predicted amino acid sequence encoded by SEQ ID NO: 533.

SEQ ID NO: 535 is a first full-length cDNA sequence for P510S.

SEQ ID NO: 536 is a second full-length cDNA sequence for P510S.

SEQ ID NO: 537 is the predicted amino acid sequence encoded by SEQ ID NO: 535.

SEQ ID NO: 538 is the predicted amino acid sequence encoded by SEQ ID NO: 536.

SEQ ID NO: 539 is the peptide P501S-370.

SEQ ID NO: 540 is the peptide P501S-376.

SEQ ID NO: 541-550 are epitopes of P501S.

SEQ ID NO: 551 corresponds to amino acids 543-553 of P501S.

DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to compositions and methods for the therapy and diagnosis of cancer, such as prostate cancer. The compositions described herein may include prostate-specific polypeptides, polynucleotides encoding such polypeptides, binding agents such as antibodies, antigen presenting cells (APCs) and/or immune system cells (*e.g.*, T cells). Polypeptides of the present invention generally comprise at least a portion (such as an immunogenic portion) of a prostate-specific protein or a variant thereof. A "prostate-specific protein" is a protein that is expressed in normal prostate and/or prostate tumor cells at a level that is at least two-fold, and preferably at least five-fold, greater than the level of expression in a non-prostate normal tissue, as determined using a representative assay provided herein. Certain prostate-specific proteins are proteins that react detectably (within an immunoassay, such as an ELISA or Western blot) with antisera of a patient afflicted with prostate cancer. Polynucleotides of the subject invention generally comprise a DNA or RNA sequence that encodes all or a portion of such a polypeptide, or that is complementary to such a sequence. Antibodies are generally immune system proteins, or antigen-binding fragments thereof, that are capable of binding to a polypeptide as described above. Antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B-cells that express a polypeptide as described above. T cells that may be employed within such compositions are generally T cells that are specific for a polypeptide as described above.

The present invention is based on the discovery of human prostate-specific proteins. Sequences of polynucleotides encoding certain prostate-specific proteins, or portions thereof, are provided in SEQ ID NOs:1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382, 384-476, 524, 526, 530, 531, 533, 535 and 536. Sequences of polypeptides comprising at least a portion of a prostate-specific protein are provided in SEQ ID NOs:112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534 and 537-550.

PROSTATE-SPECIFIC PROTEIN POLYNUCLEOTIDES

Any polynucleotide that encodes a prostate-specific protein or a portion or other variant thereof as described herein is encompassed by the present invention. Preferred

polynucleotides comprise at least 15 consecutive nucleotides, preferably at least 30 consecutive nucleotides and more preferably at least 45 consecutive nucleotides, that encode a portion of a prostate-specific protein. More preferably, a polynucleotide encodes an immunogenic portion of a prostate-specific protein. Polynucleotides complementary to any such sequences are also encompassed by the present invention. Polynucleotides may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a prostate-specific protein or a portion thereof) or may comprise a variant of such a sequence. Polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the immunogenicity of the encoded polypeptide is not diminished, relative to a native protein. The effect on the immunogenicity of the encoded polypeptide may generally be assessed as described herein. Variants preferably exhibit at least about 70% identity, more preferably at least about 80% identity and most preferably at least about 90% identity to a polynucleotide sequence that encodes a native prostate-specific protein or a portion thereof. The term "variants" also encompasses homologous genes of xenogenic origin.

Two polynucleotide or polypeptide sequences are said to be "identical" if the sequence of nucleotides or amino acids in the two sequences is the same when aligned for maximum correspondence as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices

for detecting distant relationships. In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenes pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS 5*:151-153; Myers, E.W. and Muller W. (1988) *CABIOS 4*:11-17; Robinson, E.D. (1971) *Comb. Theor 11*:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol. 4*:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA 80*:726-730.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid bases or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Variants may also, or alternatively, be substantially homologous to a native gene, or a portion or complement thereof. Such polynucleotide variants are capable of hybridizing under moderately stringent conditions to a naturally occurring DNA sequence encoding a native prostate-specific protein (or a complementary sequence). Suitable moderately stringent conditions include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-65°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS.

It will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such

as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

Polynucleotides may be prepared using any of a variety of techniques. For example, a polynucleotide may be identified, as described in more detail below, by screening a microarray of cDNAs for tumor-associated expression (*i.e.*, expression that is at least five fold greater in a prostate-specific than in normal tissue, as determined using a representative assay provided herein). Such screens may be performed using a Synteni microarray (Palo Alto, CA) according to the manufacturer's instructions (and essentially as described by Schena et al., *Proc. Natl. Acad. Sci. USA 93*:10614-10619, 1996 and Heller et al., *Proc. Natl. Acad. Sci. USA 94*:2150-2155, 1997). Alternatively, polypeptides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as prostate-specific cells. Such polynucleotides may be amplified via polymerase chain reaction (PCR). For this approach, sequence-specific primers may be designed based on the sequences provided herein, and may be purchased or synthesized.

An amplified portion may be used to isolate a full length gene from a suitable library (*e.g.*, a prostate-specific cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (*e.g.*, by nick-translation or end-labeling with ^{32}P) using well known techniques. A bacterial or bacteriophage library is then screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (*see* Sambrook et al., *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences are then assembled into

a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments; using well known techniques.

Alternatively, there are numerous amplification techniques for obtaining a full length coding sequence from a partial cDNA sequence. Within such techniques, amplification is generally performed via PCR. Any of a variety of commercially available kits may be used to perform the amplification step. Primers may be designed using, for example, software well known in the art. Primers are preferably 22-30 nucleotides in length, have a GC content of at least 50% and anneal to the target sequence at temperatures of about 68°C to 72°C. The amplified region may be sequenced as described above, and overlapping sequences assembled into a contiguous sequence.

One such amplification technique is inverse PCR (*see* Triglia et al., *Nucl. Acids Res. 16*:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom et al., *PCR Methods Applic. 1*:111-19, 1991) and walking PCR (Parker et al., *Nucl. Acids. Res. 19*:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (*e.g.*, NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences may also be obtained by analysis of genomic fragments.

Certain nucleic acid sequences of cDNA molecules encoding at least a portion of a prostate-specific protein are provided in SEQ ID NO:1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382, 384-476, 524, 526, 530, 531, 533, 535 and 536.

Isolation of these polynucleotides is described below. Each of these prostate-specific proteins was overexpressed in prostate tumor tissue.

Polynucleotide variants may generally be prepared by any method known in the art, including chemical synthesis by, for example, solid phase phosphoramidite chemical synthesis. Modifications in a polynucleotide sequence may also be introduced using standard mutagenesis techniques, such as oligonucleotide-directed site-specific mutagenesis (*see* Adelman et al., *DNA* 2:183, 1983). Alternatively, RNA molecules may be generated by *in vitro* or *in vivo* transcription of DNA sequences encoding a prostate-specific protein, or portion thereof, provided that the DNA is incorporated into a vector with a suitable RNA polymerase promoter (such as T7 or SP6). Certain portions may be used to prepare an encoded polypeptide, as described herein. In addition, or alternatively, a portion may be administered to a patient such that the encoded polypeptide is generated *in vivo* (*e.g.*, by transfecting antigen-presenting cells, such as dendritic cells, with a cDNA construct encoding a prostate-specific polypeptide, and administering the transfected cells to the patient).

A portion of a sequence complementary to a coding sequence (*i.e.*, an antisense polynucleotide) may also be used as a probe or to modulate gene expression. cDNA constructs that can be transcribed into antisense RNA may also be introduced into cells of tissues to facilitate the production of antisense RNA. An antisense polynucleotide may be used, as described herein, to inhibit expression of a protein. Antisense technology can be used to control gene expression through triple-helix formation, which compromises the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors or regulatory molecules (*see* Gee et al., *In* Huber and Carr, *Molecular and Immunologic Approaches*, Futura Publishing Co. (Mt. Kisco, NY; 1994)). Alternatively, an antisense molecule may be designed to hybridize with a control region of a gene (*e.g.*, promoter, enhancer or transcription initiation site), and block transcription of the gene; or to block translation by inhibiting binding of a transcript to ribosomes.

A portion of a coding sequence, or of a complementary sequence, may also be designed as a probe or primer to detect gene expression. Probes may be labeled with a variety of reporter groups, such as radionuclides and enzymes, and are preferably at least 10 nucleotides in length, more preferably at least 20 nucleotides in length and still more preferably at least 30 nucleotides in length. Primers, as noted above, are preferably 22-30 nucleotides in length.

Any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3'

ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl- methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and uridine.

Nucleotide sequences as described herein may be joined to a variety of other nucleotide sequences using established recombinant DNA techniques. For example, a polynucleotide may be cloned into any of a variety of cloning vectors, including plasmids, phagemids, lambda phage derivatives and cosmids. Vectors of particular interest include expression vectors, replication vectors, probe generation vectors and sequencing vectors. In general, a vector will contain an origin of replication functional in at least one organism, convenient restriction endonuclease sites and one or more selectable markers. Other elements will depend upon the desired use, and will be apparent to those of ordinary skill in the art.

Within certain embodiments, polynucleotides may be formulated so as to permit entry into a cell of a mammal, and expression therein. Such formulations are particularly useful for therapeutic purposes, as described below. Those of ordinary skill in the art will appreciate that there are many ways to achieve expression of a polynucleotide in a target cell, and any suitable method may be employed. For example, a polynucleotide may be incorporated into a viral vector such as, but not limited to, adenovirus, adeno-associated virus, retrovirus, or vaccinia or other pox virus (*e.g.*, avian pox virus). The polynucleotides may also be administered as naked plasmid vectors. Techniques for incorporating DNA into such vectors are well known to those of ordinary skill in the art. A retroviral vector may additionally transfer or incorporate a gene for a selectable marker (to aid in the identification or selection of transduced cells) and/or a targeting moiety, such as a gene that encodes a ligand for a receptor on a specific target cell, to render the vector target specific. Targeting may also be accomplished using an antibody, by methods known to those of ordinary skill in the art.

Other formulations for therapeutic purposes include colloidal dispersion systems, such as macromolecule complexes, nanocapsules, microspheres, beads, and lipid-based systems including oil-in-water emulsions, micelles, mixed micelles, and liposomes. A preferred colloidal system for use as a delivery vehicle *in vitro* and *in vivo* is a liposome (*i.e.*, an artificial membrane vesicle). The preparation and use of such systems is well known in the art.

PROSTATE-SPECIFIC POLYPEPTIDES

Within the context of the present invention, polypeptides may comprise at least an immunogenic portion of a prostate-specific protein or a variant thereof, as described herein. As noted above, a "prostate-specific protein" is a protein that is expressed by normal prostate and/or prostate tumor cells. Proteins that are prostate-specific proteins also react detectably within an immunoassay (such as an ELISA) with antisera from a patient with prostate cancer. Polypeptides as described herein may be of any length. Additional sequences derived from the native protein and/or heterologous sequences may be present, and such sequences may (but need not) possess further immunogenic or antigenic properties.

An "immunogenic portion," as used herein is a portion of a protein that is recognized (*i.e.*, specifically bound) by a B-cell and/or T-cell surface antigen receptor. Such immunogenic portions generally comprise at least 5 amino acid residues, more preferably at least 10, and still more preferably at least 20 amino acid residues of a prostate-specific protein or a variant thereof. Certain preferred immunogenic portions include peptides in which an N-terminal leader sequence and/or transmembrane domain have been deleted. Other preferred immunogenic portions may contain a small N- and/or C-terminal deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (*i.e.*, they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and antibodies may be prepared as described herein, and using well known techniques. An immunogenic portion of a native prostate-specific protein is a portion that reacts with such antisera and/or T-cells at a level that is not substantially less than the reactivity of the full length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Such immunogenic portions may react within such assays at a level that is similar to or greater than the reactivity of the full length polypeptide. Such screens may generally be performed using methods well known to those of ordinary skill in the art, such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual,* Cold Spring Harbor Laboratory, 1988. For example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the

immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, ^{125}I-labeled Protein A.

As noted above, a composition may comprise a variant of a native prostate-specific protein. A polypeptide "variant," as used herein, is a polypeptide that differs from a native prostate-specific protein in one or more substitutions, deletions, additions and/or insertions, such that the immunogenicity of the polypeptide is not substantially diminished. In other words, the ability of a variant to react with antigen-specific antisera may be enhanced or unchanged, relative to the native protein, or may be diminished by less than 50%, and preferably less than 20%, relative to the native protein. Such variants may generally be identified by modifying one of the above polypeptide sequences and evaluating the reactivity of the modified polypeptide with antigen-specific antibodies or antisera as described herein. Preferred variants include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other preferred variants include variants in which a small portion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein. Polypeptide variants preferably exhibit at least about 70%, more preferably at least about 90% and most preferably at least about 95% identity (determined as described above) to the identified polypeptides.

Preferably, a variant contains conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. Amino acid substitutions may generally be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino

acids that have minimal influence on the immunogenicity, secondary structure and hydropathic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (*e.g.*, poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

Polypeptides may be prepared using any of a variety of well known techniques. Recombinant polypeptides encoded by DNA sequences as described above may be readily prepared from the DNA sequences using any of a variety of expression vectors known to those of ordinary skill in the art. Expression may be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide. Suitable host cells include prokaryotes, yeast, higher eukaryotic and plant cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line such as COS or CHO. Supernatants from suitable host/vector systems which secrete recombinant protein or polypeptide into culture media may be first concentrated using a commercially available filter. Following concentration, the concentrate may be applied to a suitable purification matrix such as an affinity matrix or an ion exchange resin. Finally, one or more reverse phase HPLC steps can be employed to further purify a recombinant polypeptide.

Portions and other variants having fewer than about 100 amino acids, and generally fewer than about 50 amino acids, may also be generated by synthetic means, using techniques well known to those of ordinary skill in the art. For example, such polypeptides may be synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. *See* Merrifield, *J. Am. Chem. Soc. 85*:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

Within certain specific embodiments, a polypeptide may be a fusion protein that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known prostate-specific protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner),

preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the protein or to enable the protein to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the protein.

Fusion proteins may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion protein is expressed as a recombinant protein, allowing the production of increased levels, relative to a non-fused protein, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion protein that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and the second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion protein using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene 40*:39-46, 1985; Murphy et al., *Proc. Natl. Acad. Sci. USA 83*:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are

located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

Fusion proteins are also provided that comprise a polypeptide of the present invention together with an unrelated immunogenic protein. Preferably the immunogenic protein is capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (*see*, for example, Stoute et al. *New Engl. J. Med.*, *336*:86-91, 1997).

Within preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium Haemophilus influenza B (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third of the protein (*e.g.*, the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemaglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the LytA gene; *Gene 43*:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (*see Biotechnology 10*:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion protein. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

In general, polypeptides (including fusion proteins) and polynucleotides as described herein are isolated. An "isolated" polypeptide or polynucleotide is one that is removed from its

original environment. For example, a naturally-occurring protein is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure. A polynucleotide is considered to be isolated if, for example, it is cloned into a vector that is not a part of the natural environment.

BINDING AGENTS

The present invention further provides agents, such as antibodies and antigen-binding fragments thereof, that specifically bind to a prostate-specific protein. As used herein, an antibody, or antigen-binding fragment thereof, is said to "specifically bind" to a prostate-specific protein if it reacts at a detectable level (within, for example, an ELISA) with a prostate-specific protein, and does not react detectably with unrelated proteins under similar conditions. As used herein, "binding" refers to a noncovalent association between two separate molecules such that a complex is formed. The ability to bind may be evaluated by, for example, determining a binding constant for the formation of the complex. The binding constant is the value obtained when the concentration of the complex is divided by the product of the component concentrations. In general, two compounds are said to "bind," in the context of the present invention, when the binding constant for complex formation exceeds about 10^3 L/mol. The binding constant may be determined using methods well known in the art.

Binding agents may be further capable of differentiating between patients with and without a cancer, such as prostate cancer, using the representative assays provided herein. In other words, antibodies or other binding agents that bind to a prostate-specific protein will generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, and will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*, blood, sera, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. It will be apparent that a statistically significant number of samples with and without the disease should be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Most preferably, antibodies employed in the inventive methods have the ability to induce lysis of tumor cells by activation of complement and mediation of antibody-dependent cellular cytotoxicity (ADCC). Antibodies of different classes and subclasses differ in these properties. For example, mouse antibodies of the IgG2a and IgG3 classes are capable of activating serum complement upon binding to target cells which express the antigen against which the antibodies were raised, and can mediate ADCC.

Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.*, mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol. 6*:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells

and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

The preparation of mouse and rabbit monoclonal antibodies that specifically bind to polypeptides of the present invention is described in detail below. However, the antibodies of the present invention are not limited to those derived from mice. Human antibodies may also be employed in the inventive methods and may prove to be preferable. Such antibodies can be obtained using human hybridomas as described by Cote *et al.* (Monoclonal Antibodies and Cancer Therapy, Alan R. Lisa, p. 77, 1985). The present invention also encompasses antibodies made by recombinant means such as chimeric antibodies, wherein the variable region and constant region are derived from different species, and CDR-grafted antibodies, wherein the complementarity determining region is derived from a different species, as described in US Patents 4,816,567 and 5,225,539. Chimeric antibodies may be prepared by splicing genes for a mouse antibody molecule having a desired antigen specificity together with genes for a human antibody molecule having the desired biological activity, such as activation of human complement and mediation of ADCC (Morrison *et al. Proc. Natl. Acad. Sci. USA 81*:6851, 1984; Neuberger *et al. Nature 312*:604, 1984; Takeda *et al. Nature 314*:452, 1985).

Within certain embodiments, the use of antigen-binding fragments of antibodies may be preferred. Such fragments include Fab fragments, which may be prepared using standard techniques. Briefly, immunoglobulins may be purified from rabbit serum by affinity chromatography on Protein A bead columns (Harlow and Lane, *Antibodies: A Laboratory Manual*,

Cold Spring Harbor Laboratory, 1988) and digested by papain to yield Fab and Fc fragments. The Fab and Fc fragments may be separated by affinity chromatography on protein A bead columns.

Monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ^{90}Y, ^{123}I, ^{125}I, ^{131}I, ^{186}Re, ^{188}Re, ^{211}At, and ^{212}Bi. Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diptheria toxin, cholera toxin, gelonin, Pseudomonas exotoxin, Shigella toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to

Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn et al.), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler et al.).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers which provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating compounds and their synthesis.

A variety of routes of administration for the antibodies and immunoconjugates may be used. Typically, administration will be intravenous, intramuscular, subcutaneous or in the bed of a resected tumor. It will be evident that the precise dose of the antibody/immunoconjugate will vary depending upon the antibody used, the antigen density on the tumor, and the rate of clearance of the antibody.

T CELLS

Immunotherapeutic compositions may also, or alternatively, comprise T cells specific for a prostate-specific protein. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral

blood, or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the ISOLEX™ system, available from Nexell Therapeutics Inc., Irvine, CA (see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with a prostate-specific polypeptide, polynucleotide encoding a prostate-specific polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide. Preferably, a prostate-specific polypeptide or polynucleotide is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a prostate-specific polypeptide if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen et al., *Cancer Res.* *54*:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (*e.g.*, by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a prostate-specific polypeptide (100 ng/ml - 100 μg/ml, preferably 200 ng/ml - 25 μg/ml) for 3 - 7 days should result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (*e.g.*, TNF or IFN-γ) is indicative of T cell activation (*see* Coligan et al., Current Protocols in Immunology, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a prostate-specific polypeptide, polynucleotide or polypeptide-expressing APC may be $CD4^+$ and/or $CD8^+$. Prostate-specific protein-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from either a patient or a related, or unrelated, donor and are administered to the patient following stimulation and expansion.

For therapeutic purposes, $CD4^+$ or $CD8^+$ T cells that proliferate in response to a prostate-specific polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a prostate-specific polypeptide, or a short peptide corresponding to an immunogenic portion of such a polypeptide, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a prostate-specific polypeptide. Alternatively, one or more T cells that proliferate in the presence of a prostate-specific protein can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

Pharmaceutical Compositions and Vaccines

Within certain aspects, polypeptides, polynucleotides, T cells and/or binding agents disclosed herein may be incorporated into pharmaceutical compositions or immunogenic compositions (*i.e.,* vaccines). Pharmaceutical compositions comprise one or more such compounds and a physiologically acceptable carrier. Vaccines may comprise one or more such compounds and an immunostimulant. An immunostimulant may be any substance that enhances an immune response to an exogenous antigen. Examples of immunostimulants include adjuvants, biodegradable microspheres (*e.g.,* polylactic galactide) and liposomes (into which the compound is incorporated; *see e.g.,* Fullerton, U.S. Patent No. 4,235,877). Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Pharmaceutical compositions and vaccines within the scope of the present invention may also contain other compounds, which may be biologically active or inactive. For example, one or more immunogenic portions of other tumor antigens may be present, either incorporated into a fusion polypeptide or as a separate compound, within the composition or vaccine.

A pharmaceutical composition or vaccine may contain DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, bacteria and viral expression systems. Numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems 15*:143-198, 1998, and references cited therein. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression

in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope. In a preferred embodiment, the DNA may be introduced using a viral expression system (*e.g.*, vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are disclosed, for example, in Fisher-Hoch et al., *Proc. Natl. Acad. Sci. USA 86*:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci. 569*:86-103, 1989; Flexner et al., *Vaccine 8*:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques 6*:616-627, 1988; Rosenfeld et al., *Science 252*:431-434, 1991; Kolls et al., *Proc. Natl. Acad. Sci. USA 91*:215-219, 1994; Kass-Eisler et al., *Proc. Natl. Acad. Sci. USA 90*:11498-11502, 1993; Guzman et al., *Circulation 88*:2838-2848, 1993; and Guzman et al., *Cir. Res. 73*:1202-1207, 1993. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer et al., *Science 259*:1745-1749, 1993 and reviewed by Cohen, *Science 259*:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of carrier will vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (*e.g.*, polylactate polyglycolate) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268 and 5,075,109.

Such compositions may also comprise buffers (*e.g.*, neutral buffered saline or phosphate buffered saline), carbohydrates (*e.g.*, glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, chelating agents such as EDTA

or glutathione, adjuvants (*e.g.*, aluminum hydroxide) and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate. Compounds may also be encapsulated within liposomes using well known technology.

Any of a variety of immunostimulants may be employed in the vaccines of this invention. For example, an adjuvant may be included. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bortadella pertussis* or *Mycobacterium tuberculosis* derived proteins. Suitable adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, may also be used as adjuvants.

Within the vaccines provided herein, the adjuvant composition is preferably designed to induce an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (*e.g.*, IFN-γ, TNFα, IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.*, IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol.* 7:145-173, 1989.

Preferred adjuvants for use in eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A (3D-MPL), together with an aluminum salt. MPL adjuvants are available from Ribi ImmunoChem Research Inc. (Hamilton, MT; *see* US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555. Another preferred adjuvant is a saponin, preferably QS21, which may be used alone or in combination with other adjuvants. For example,

an enhanced system involves the combination of a monophosphoryl lipid A and saponin derivative, such as the combination of QS21 and 3D-MPL as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprises an oil-in-water emulsion and tocopherol. A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil-in-water emulsion is described in WO 95/17210. Any vaccine provided herein may be prepared using well known methods that result in a combination of antigen, immune response enhancer and a suitable carrier or excipient.

The compositions described herein may be administered as part of a sustained release formulation (*i.e.*, a formulation such as a capsule, sponge or gel (composed of polysaccharides for example) that effects a slow release of compound following administration). Such formulations may generally be prepared using well known technology and administered by, for example, oral, rectal or subcutaneous implantation, or by implantation at the desired target site. Sustained-release formulations may contain a polypeptide, polynucleotide or antibody dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane. Carriers for use within such formulations are biocompatible, and may also be biodegradable; preferably the formulation provides a relatively constant level of active component release. The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

Any of a variety of delivery vehicles may be employed within pharmaceutical compositions and vaccines to facilitate production of an antigen-specific immune response that targets tumor cells. Delivery vehicles include antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature 392*:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman and Levy,

Ann. Rev. Med. 50:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take-up, process and present antigens with high efficiency, and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (*see* Zitvogel et al., *Nature Med. 4*:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNFα to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNFα, CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fcγ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (*e.g.*, CD54 and CD11) and costimulatory molecules (*e.g.*, CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide encoding a prostate-specific protein (or portion or other variant thereof) such that the prostate-specific polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a composition or vaccine comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection

that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi et al., *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the prostate-specific polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (*e.g.*, vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (*e.g.*, a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

CANCER THERAPY

In further aspects of the present invention, the compositions described herein may be used for immunotherapy of cancer, such as prostate cancer. Within such methods, pharmaceutical compositions and vaccines are typically administered to a patient. As used herein, a "patient" refers to any warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions and vaccines may be used to prevent the development of a cancer or to treat a patient afflicted with a cancer. A cancer may be diagnosed using criteria generally accepted in the art, including the presence of a malignant tumor. Pharmaceutical compositions and vaccines may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides disclosed herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as $CD8^+$ cytotoxic T lymphocytes and $CD4^+$ T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer

cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow in vivo and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (*see*, for example, Cheever et al., *Immunological Reviews 157*:177, 1997).

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions disclosed herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and vaccines may be administered by injection (*e.g.*, intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (*e.g.*, by aspiration) or orally. Preferably, between 1 and 10 doses may be administered

over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 25 µg to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a prostate-specific protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

METHODS FOR DETECTING CANCER

In general, a cancer may be detected in a patient based on the presence of one or more prostate-specific proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as prostate cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer.

In general, a prostate tumor sequence should be present at a level that is at least three fold higher in tumor tissue than in normal tissue

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. *See*, *e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual,* Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length prostate-specific proteins and portions thereof to which the binding agent binds, as described above.

The solid support may be any material known to those of ordinary skill in the art to which the protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a

membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μg, and preferably about 100 ng to about 1 μg, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding partner (*see*, *e.g.*, Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with prostate cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by

assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as prostate cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along

the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1μg, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to those of ordinary skill in the art that the above protocols may be readily modified to use prostate-specific polypeptides to detect antibodies that bind to such polypeptides in a biological sample. The detection of such prostate-specific protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with a prostate-specific protein in a biological sample. Within certain methods, a biological sample comprising $CD4^+$ and/or $CD8^+$ T cells isolated from a patient is incubated with a prostate-specific polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that

expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with prostate-specific polypeptide (*e.g.*, 5 - 25 μg/ml). It may be desirable to incubate another aliquot of a T cell sample in the absence of prostate-specific polypeptide to serve as a control. For $CD4^+$ T cells, activation is preferably detected by evaluating proliferation of the T cells. For $CD8^+$ T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a prostate-specific protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a prostate-specific cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the prostate-specific protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a prostate-specific protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a prostate-specific protein that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes will hybridize to a polynucleotide encoding a polypeptide disclosed herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382, 384-476, 524, 526, 530, 531, 533, 535 and 536. Techniques for both PCR based assays and hybridization assays

are well known in the art (*see*, for example, Mullis et al., *Cold Spring Harbor Symp. Quant. Biol.*, *51*:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

In another embodiment, the disclosed compositions may be used as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

As noted above, to improve sensitivity, multiple prostate-specific protein markers may be assayed within a given sample. It will be apparent that binding agents specific for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of protein markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for proteins provided herein may be combined with assays for other known tumor antigens.

DIAGNOSTIC KITS

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a prostate-specific protein. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a prostate-specific protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a prostate-specific protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a prostate-specific protein.

The following Examples are offered by way of illustration and not by way of limitation.

EXAMPLES

EXAMPLE 1

ISOLATION AND CHARACTERIZATION OF PROSTATE-SPECIFIC POLYPEPTIDES

This Example describes the isolation of certain prostate-specific polypeptides from a prostate tumor cDNA library.

A human prostate tumor cDNA expression library was constructed from prostate tumor poly A^+ RNA using a Superscript Plasmid System for cDNA Synthesis and Plasmid Cloning kit (BRL Life Technologies, Gaithersburg, MD 20897) following the manufacturer's protocol. Specifically, prostate tumor tissues were homogenized with polytron (Kinematica, Switzerland) and total RNA was extracted using Trizol reagent (BRL Life Technologies) as directed by the manufacturer. The poly A^+ RNA was then purified using a Qiagen oligotex spin column mRNA purification kit (Qiagen, Santa Clarita, CA 91355) according to the manufacturer's protocol. First-strand cDNA was synthesized using the NotI/Oligo-dT18 primer. Double-stranded cDNA was synthesized, ligated with EcoRI/BAXI adaptors (Invitrogen, San Diego, CA) and digested with NotI. Following size fractionation with Chroma Spin-1000 columns (Clontech, Palo Alto, CA), the cDNA was ligated into the EcoRI/NotI site of pCDNA3.1 (Invitrogen) and transformed into ElectroMax *E. coli* DH10B cells (BRL Life Technologies) by electroporation.

Using the same procedure, a normal human pancreas cDNA expression library was prepared from a pool of six tissue specimens (Clontech). The cDNA libraries were characterized by determining the number of independent colonies, the percentage of clones that carried insert, the average insert size and by sequence analysis. The prostate tumor library contained 1.64×10^7 independent colonies, with 70% of clones having an insert and the average insert size being 1745 base pairs. The normal pancreas cDNA library contained 3.3×10^6 independent colonies, with 69% of clones having inserts and the average insert size being 1120 base pairs. For both libraries, sequence analysis showed that the majority of clones had a full length cDNA sequence and were synthesized from mRNA, with minimal rRNA and mitochondrial DNA contamination.

cDNA library subtraction was performed using the above prostate tumor and normal pancreas cDNA libraries, as described by Hara *et al.* (*Blood, 84*:189-199, 1994) with some modifications. Specifically, a prostate tumor-specific subtracted cDNA library was generated as

follows. Normal pancreas cDNA library (70 μg) was digested with EcoRI, NotI, and SfuI, followed by a filling-in reaction with DNA polymerase Klenow fragment. After phenol-chloroform extraction and ethanol precipitation, the DNA was dissolved in 100 μl of H_2O, heat-denatured and mixed with 100 μl (100 μg) of Photoprobe biotin (Vector Laboratories, Burlingame, CA). As recommended by the manufacturer, the resulting mixture was irradiated with a 270 W sunlamp on ice for 20 minutes. Additional Photoprobe biotin (50 μl) was added and the biotinylation reaction was repeated. After extraction with butanol five times, the DNA was ethanol-precipitated and dissolved in 23 μl H_2O to form the driver DNA.

To form the tracer DNA, 10 μg prostate tumor cDNA library was digested with BamHI and XhoI, phenol chloroform extracted and passed through Chroma spin-400 columns (Clontech). Following ethanol precipitation, the tracer DNA was dissolved in 5 μl H_2O. Tracer DNA was mixed with 15 μl driver DNA and 20 μl of 2 x hybridization buffer (1.5 M NaCl/10 mM EDTA/50 mM HEPES pH 7.5/0.2% sodium dodecyl sulfate), overlaid with mineral oil, and heat-denatured completely. The sample was immediately transferred into a 68 ^{0}C water bath and incubated for 20 hours (long hybridization [LH]). The reaction mixture was then subjected to a streptavidin treatment followed by phenol/chloroform extraction. This process was repeated three more times. Subtracted DNA was precipitated, dissolved in 12 μl H_2O, mixed with 8 μl driver DNA and 20 μl of 2 x hybridization buffer, and subjected to a hybridization at 68 ^{0}C for 2 hours (short hybridization [SH]). After removal of biotinylated double-stranded DNA, subtracted cDNA was ligated into BamHI/XhoI site of chloramphenicol resistant pBCSK$^+$ (Stratagene, La Jolla, CA 92037) and transformed into ElectroMax *E. coli* DH10B cells by electroporation to generate a prostate tumor specific subtracted cDNA library (referred to as "prostate subtraction 1").

To analyze the subtracted cDNA library, plasmid DNA was prepared from 100 independent clones, randomly picked from the subtracted prostate tumor specific library and grouped based on insert size. Representative cDNA clones were further characterized by DNA sequencing with a Perkin Elmer/Applied Biosystems Division Automated Sequencer Model 373A (Foster City, CA). Six cDNA clones, hereinafter referred to as F1-13, F1-12, F1-16, H1-1, H1-9 and H1-4, were shown to be abundant in the subtracted prostate-specific cDNA library. The determined 3' and 5' cDNA sequences for F1-12 are provided in SEQ ID NO: 2 and 3, respectively, with determined 3' cDNA sequences for F1-13, F1-16, H1-1, H1-9 and H1-4 being provided in SEQ ID NO: 1 and 4-7, respectively.

The cDNA sequences for the isolated clones were compared to known sequences in the gene bank using the EMBL and GenBank databases (release 96). Four of the prostate tumor cDNA clones, F1-13, F1-16, H1-1, and H1-4, were determined to encode the following previously identified proteins: prostate specific antigen (PSA), human glandular kallikrein, human tumor expression enhanced gene, and mitochondria cytochrome C oxidase subunit II. H1-9 was found to be identical to a previously identified human autonomously replicating sequence. No significant homologies to the cDNA sequence for F1-12 were found.

Subsequent studies led to the isolation of a full-length cDNA sequence for F1-12. This sequence is provided in SEQ ID NO: 107, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 108.

To clone less abundant prostate tumor specific genes, cDNA library subtraction was performed by subtracting the prostate tumor cDNA library described above with the normal pancreas cDNA library and with the three most abundant genes in the previously subtracted prostate tumor specific cDNA library: human glandular kallikrein, prostate specific antigen (PSA), and mitochondria cytochrome C oxidase subunit II. Specifically, 1 μg each of human glandular kallikrein, PSA and mitochondria cytochrome C oxidase subunit II cDNAs in pCDNA3.1 were added to the driver DNA and subtraction was performed as described above to provide a second subtracted cDNA library hereinafter referred to as the "subtracted prostate tumor specific cDNA library with spike".

Twenty-two cDNA clones were isolated from the subtracted prostate tumor specific cDNA library with spike. The determined 3' and 5' cDNA sequences for the clones referred to as J1-17, L1-12, N1-1862, J1-13, J1-19, J1-25, J1-24, K1-58, K1-63, L1-4 and L1-14 are provided in SEQ ID NOS: 8-9, 10-11, 12-13, 14-15, 16-17, 18-19, 20-21, 22-23, 24-25, 26-27 and 28-29, respectively. The determined 3' cDNA sequences for the clones referred to as J1-12, J1-16, J1-21, K1-48, K1-55, L1-2, L1-6, N1-1858, N1-1860, N1-1861, N1-1864 are provided in SEQ ID NOS: 30-40, respectively. Comparison of these sequences with those in the gene bank as described above, revealed no significant homologies to three of the five most abundant DNA species, (J1-17, L1-12 and N1-1862; SEQ ID NOS: 8-9, 10-11 and 12-13, respectively). Of the remaining two most abundant species, one (J1-12; SEQ ID NO:30) was found to be identical to the previously identified human pulmonary surfactant-associated protein, and the other (K1-48; SEQ ID NO:33) was determined to have some homology to *R. norvegicus* mRNA for 2-arylpropionyl-CoA epimerase. Of the 17 less abundant cDNA clones isolated from the subtracted prostate tumor specific cDNA

library with spike, four (J1-16, K1-55, L1-6 and N1-1864; SEQ ID NOS:31, 34, 36 and 40, respectively) were found to be identical to previously identified sequences, two (J1-21 and N1-1860; SEQ ID NOS: 32 and 38, respectively) were found to show some homology to non-human sequences, and two (L1-2 and N1-1861; SEQ ID NOS: 35 and 39, respectively) were found to show some homology to known human sequences. No significant homologies were found to the polypeptides J1-13, J1-19, J1-24, J1-25, K1-58, K1-63, L1-4, L1-14 (SEQ ID NOS: 14-15, 16-17, 20-21, 18-19, 22-23, 24-25, 26-27, 28-29, respectively).

Subsequent studies led to the isolation of full length cDNA sequences for J1-17, L1-12 and N1-1862 (SEQ ID NOS: 109-111, respectively). The corresponding predicted amino acid sequences are provided in SEQ ID NOS: 112-114. L1-12 is also referred to as P501S.

In a further experiment, four additional clones were identified by subtracting a prostate tumor cDNA library with normal prostate cDNA prepared from a pool of three normal prostate poly A+ RNA (referred to as "prostate subtraction 2"). The determined cDNA sequences for these clones, hereinafter referred to as U1-3064, U1-3065, V1-3692 and 1A-3905, are provided in SEQ ID NO: 69-72, respectively. Comparison of the determined sequences with those in the gene bank revealed no significant homologies to U1-3065.

A second subtraction with spike (referred to as "prostate subtraction spike 2") was performed by subtracting a prostate tumor specific cDNA library with spike with normal pancreas cDNA library and further spiked with PSA, J1-17, pulmonary surfactant-associated protein, mitochondrial DNA, cytochrome c oxidase subunit II, N1-1862, autonomously replicating sequence, L1-12 and tumor expression enhanced gene. Four additional clones, hereinafter referred to as V1-3686, R1-2330, 1B-3976 and V1-3679, were isolated. The determined cDNA sequences for these clones are provided in SEQ ID NO:73-76, respectively. Comparison of these sequences with those in the gene bank revealed no significant homologies to V1-3686 and R1-2330.

Further analysis of the three prostate subtractions described above (prostate subtraction 2, subtracted prostate tumor specific cDNA library with spike, and prostate subtraction spike 2) resulted in the identification of sixteen additional clones, referred to as 1G-4736, 1G-4738, 1G-4741, 1G-4744, 1G-4734, 1H-4774, 1H-4781, 1H-4785, 1H-4787, 1H-4796, 1I-4810, 1I-4811, 1J-4876, 1K-4884 and 1K-4896. The determined cDNA sequences for these clones are provided in SEQ ID NOS: 77-92, respectively. Comparison of these sequences with those in the gene bank as described above, revealed no significant homologies to 1G-4741, 1G-4734, 1I-4807, 1J-4876 and 1K-4896 (SEQ ID NOS: 79, 81, 87, 90 and 92, respectively). Further analysis of the isolated

clones led to the determination of extended cDNA sequences for 1G-4736, 1G-4738, 1G-4741, 1G-4744, 1H-4774, 1H-4781, 1H-4785, 1H-4787, 1H-4796, 1I-4807, 1J-4876, 1K-4884 and 1K-4896, provided in SEQ ID NOS: 179-188 and 191-193, respectively, and to the determination of additional partial cDNA sequences for 1I-4810 and 1I-4811, provided in SEQ ID NOS: 189 and 190, respectively.

Additional studies with prostate subtraction spike 2 resulted in the isolation of three more clones. Their sequences were determined as described above and compared to the most recent GenBank. All three clones were found to have homology to known genes, which are Cysteine-rich protein, KIAA0242, and KIAA0280 (SEQ ID NO: 317, 319, and 320, respectively). Further analysis of these clones by Synteni microarray (Synteni, Palo Alto, CA) demonstrated that all three clones were over-expressed in most prostate tumors and prostate BPH, as well as in the majority of normal prostate tissues tested, but low expression in all other normal tissues.

An additional subtraction was performed by subtracting a normal prostate cDNA library with normal pancreas cDNA (referred to as "prostate subtraction 3"). This led to the identification of six additional clones referred to as 1G-4761, 1G-4762, 1H-4766, 1H-4770, 1H-4771 and 1H-4772 (SEQ ID NOS: 93-98). Comparison of these sequences with those in the gene bank revealed no significant homologies to 1G-4761 and 1H-4771 (SEQ ID NOS: 93 and 97, respectively). Further analysis of the isolated clones led to the determination of extended cDNA sequences for 1G-4761, 1G-4762, 1H-4766 and 1H-4772 provided in SEQ ID NOS: 194-196 and 199, respectively, and to the determination of additional partial cDNA sequences for 1H-4770 and 1H-4771, provided in SEQ ID NOS: 197 and 198, respectively.

Subtraction of a prostate tumor cDNA library, prepared from a pool of polyA+ RNA from three prostate cancer patients, with a normal pancreas cDNA library (prostate subtraction 4) led to the identification of eight clones, referred to as 1D-4297, 1D-4309, 1D.1-4278, 1D-4288, 1D-4283, 1D-4304, 1D-4296 and 1D-4280 (SEQ ID NOS: 99-107). These sequences were compared to those in the gene bank as described above. No significant homologies were found to 1D-4283 and 1D-4304 (SEQ ID NOS: 103 and 104, respectively). Further analysis of the isolated clones led to the determination of extended cDNA sequences for 1D-4309, 1D.1-4278, 1D-4288, 1D-4283, 1D-4304, 1D-4296 and 1D-4280, provided in SEQ ID NOS: 200-206, respectively.

cDNA clones isolated in prostate subtraction 1 and prostate subtraction 2, described above, were colony PCR amplified and their mRNA expression levels in prostate tumor, normal prostate and in various other normal tissues were determined using microarray technology (Synteni,

Palo Alto, CA). Briefly, the PCR amplification products were dotted onto slides in an array format, with each product occupying a unique location in the array. mRNA was extracted from the tissue sample to be tested, reverse transcribed, and fluorescent-labeled cDNA probes were generated. The microarrays were probed with the labeled cDNA probes, the slides scanned and fluorescence intensity was measured. This intensity correlates with the hybridization intensity. Two clones (referred to as P509S and P510S) were found to be over-expressed in prostate tumor and normal prostate and expressed at low levels in all other normal tissues tested (liver, pancreas, skin, bone marrow, brain, breast, adrenal gland, bladder, testes, salivary gland, large intestine, kidney, ovary, lung, spinal cord, skeletal muscle and colon). The determined cDNA sequences for P509S and P510S are provided in SEQ ID NO: 223 and 224, respectively. Comparison of these sequences with those in the gene bank as described above, revealed some homology to previously identified ESTs.

Additional, studies led to the isolation of the full-length cDNA sequence for P509S. This sequence is provided in SEQ ID NO: 332, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 339. Two variant full-length cDNA sequences for P510S are provided in SEQ ID NO: 535 and 536, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 537 and 538, respectively.

EXAMPLE 2

DETERMINATION OF TISSUE SPECIFICITY OF PROSTATE-SPECIFIC POLYPEPTIDES

Using gene specific primers, mRNA expression levels for the representative prostate-specific polypeptides F1-16, H1-1, J1-17 (also referred to as P502S), L1-12 (also referred to as P501S), F1-12 (also referred to as P504S) and N1-1862 (also referred to as P503S) were examined in a variety of normal and tumor tissues using RT-PCR.

Briefly, total RNA was extracted from a variety of normal and tumor tissues using Trizol reagent as described above. First strand synthesis was carried out using 1-2 μg of total RNA with SuperScript II reverse transcriptase (BRL Life Technologies) at 42 ^{0}C for one hour. The cDNA was then amplified by PCR with gene-specific primers. To ensure the semi-quantitative nature of the RT-PCR, β-actin was used as an internal control for each of the tissues examined. First, serial dilutions of the first strand cDNAs were prepared and RT-PCR assays were performed using β-actin specific primers. A dilution was then chosen that enabled the linear range amplification of the β-actin template and which was sensitive enough to reflect the differences in the initial copy numbers. Using these conditions, the β-actin levels were determined for each

reverse transcription reaction from each tissue. DNA contamination was minimized by DNase treatment and by assuring a negative PCR result when using first strand cDNA that was prepared without adding reverse transcriptase.

mRNA Expression levels were examined in four different types of tumor tissue (prostate tumor from 2 patients, breast tumor from 3 patients, colon tumor, lung tumor), and sixteen different normal tissues, including prostate, colon, kidney, liver, lung, ovary, pancreas, skeletal muscle, skin, stomach, testes, bone marrow and brain. F1-16 was found to be expressed at high levels in prostate tumor tissue, colon tumor and normal prostate, and at lower levels in normal liver, skin and testes, with expression being undetectable in the other tissues examined. H1-1 was found to be expressed at high levels in prostate tumor, lung tumor, breast tumor, normal prostate, normal colon and normal brain, at much lower levels in normal lung, pancreas, skeletal muscle, skin, small intestine, bone marrow, and was not detected in the other tissues tested. J1-17 (P502S) and L1-12 (P501S) appear to be specifically over-expressed in prostate, with both genes being expressed at high levels in prostate tumor and normal prostate but at low to undetectable levels in all the other tissues examined. N1-1862 (P503S) was found to be over-expressed in 60% of prostate tumors and detectable in normal colon and kidney. The RT-PCR results thus indicate that F1-16, H1-1, J1-17 (P502S), N1-1862 (P503S) and L1-12 (P501S) are either prostate specific or are expressed at significantly elevated levels in prostate.

Further RT-PCR studies showed that F1-12 (P504S) is over-expressed in 60% of prostate tumors, detectable in normal kidney but not detectable in all other tissues tested. Similarly, R1-2330 was shown to be over-expressed in 40% of prostate tumors, detectable in normal kidney and liver, but not detectable in all other tissues tested. U1-3064 was found to be over-expressed in 60% of prostate tumors, and also expressed in breast and colon tumors, but was not detectable in normal tissues.

RT-PCR characterization of R1-2330, U1-3064 and 1D-4279 showed that these three antigens are over-expressed in prostate and/or prostate tumors.

Northern analysis with four prostate tumors, two normal prostate samples, two BPH prostates, and normal colon, kidney, liver, lung, pancrease, skeletal muscle, brain, stomach, testes, small intestine and bone marrow, showed that L1-12 (P501S) is over-expressed in prostate tumors and normal prostate, while being undetectable in other normal tissues tested. J1-17 (P502S) was detected in two prostate tumors and not in the other tissues tested. N1-1862 (P503S) was found to be over-expressed in three prostate tumors and to be expressed in normal prostate, colon and kidney,

but not in other tissues tested. F1-12 (P504S) was found to be highly expressed in two prostate tumors and to be undetectable in all other tissues tested.

The microarray technology described above was used to determine the expression levels of representative antigens described herein in prostate tumor, breast tumor and the following normal tissues: prostate, liver, pancreas, skin, bone marrow, brain, breast, adrenal gland, bladder, testes, salivary gland, large intestine, kidney, ovary, lung, spinal cord, skeletal muscle and colon. L1-12 (P501S) was found to be over-expressed in normal prostate and prostate tumor, with some expression being detected in normal skeletal muscle. Both J1-12 and F1-12 (P504S) were found to be over-expressed in prostate tumor, with expression being lower or undetectable in all other tissues tested. N1-1862 (P503S) was found to be expressed at high levels in prostate tumor and normal prostate, and at low levels in normal large intestine and normal colon, with expression being undetectable in all other tissues tested. R1-2330 was found to be over-expressed in prostate tumor and normal prostate, and to be expressed at lower levels in all other tissues tested. 1D-4279 was found to be over-expressed in prostate tumor and normal prostate, expressed at lower levels in normal spinal cord, and to be undetectable in all other tissues tested.

Further microarray analysis to specifically address the extent to which P501S (SEQ ID NO: 110) was expressed in breast tumor revealed moderate over-expression not only in breast tumor, but also in metastatic breast tumor (2/31), with negligible to low expression in normal tissues. This data suggests that P501S may be over-expressed in various breast tumors as well as in prostate tumors.

The expression levels of 32 ESTs (expressed sequence tags) described by Vasmatzis *et al.* (*Proc. Natl. Acad. Sci. USA 95*:300-304, 1998) in a variety of tumor and normal tissues were examined by microarray technology as described above. Two of these clones (referred to as P1000C and P1001C) were found to be over-expressed in prostate tumor and normal prostate, and expressed at low to undetectable levels in all other tissues tested (normal aorta, thymus, resting and activated PBMC, epithelial cells, spinal cord, adrenal gland, fetal tissues, skin, salivary gland, large intestine, bone marrow, liver, lung, dendritic cells, stomach, lymph nodes, brain, heart, small intestine, skeletal muscle, colon and kidney. The determined cDNA sequences for P1000C and P1001C are provided in SEQ ID NO: 384 and 472, respectively. The sequence of P1001C was found to show some homology to the previously isolated Human mRNA for JM27 protein. No significant homologies were found to the sequence of P1000C.

The expression of the polypeptide encoded by the full length cDNA sequence for F1-12 (also referred to as P504S; SEQ ID NO: 108) was investigated by immunohistochemical analysis. Rabbit-anti-P504S polyclonal antibodies were generated against the full length P504S protein by standard techniques. Subsequent isolation and characterization of the polyclonal antibodies were also performed by techniques well known in the art. Immunohistochemical analysis showed that the P504S polypeptide was expressed in 100% of prostate carcinoma samples tested (n=5).

The rabbit-anti-P504S polyclonal antibody did not appear to label benign prostate cells with the same cytoplasmic granular staining, but rather with light nuclear staining. Analysis of normal tissues revealed that the encoded polypeptide was found to be expressed in some, but not all normal human tissues. Positive cytoplasmic staining with rabbit-anti-P504S polyclonal antibody was found in normal human kidney, liver, brain, colon and lung-associated macrophages, whereas heart and bone marrow were negative.

This data indicates that the P504S polypeptide is present in prostate cancer tissues, and that there are qualitative and quantitative differences in the staining between benign prostatic hyperplasia tissues and prostate cancer tissues, suggesting that this polypeptide may be detected selectively in prostate tumors and therefore be useful in the diagnosis of prostate cancer.

EXAMPLE 3

ISOLATION AND CHARACTERIZATION OF PROSTATE-SPECIFIC POLYPEPTIDES BY PCR-BASED SUBTRACTION

A cDNA subtraction library, containing cDNA from normal prostate subtracted with ten other normal tissue cDNAs (brain, heart, kidney, liver, lung, ovary, placenta, skeletal muscle, spleen and thymus) and then submitted to a first round of PCR amplification, was purchased from Clontech. This library was subjected to a second round of PCR amplification, following the manufacturer's protocol. The resulting cDNA fragments were subcloned into the vector pT7 Blue T-vector (Novagen, Madison, WI) and transformed into XL-1 Blue MRF' *E. coli* (Stratagene). DNA was isolated from independent clones and sequenced using a Perkin Elmer/Applied Biosystems Division Automated Sequencer Model 373A.

Fifty-nine positive clones were sequenced. Comparison of the DNA sequences of these clones with those in the gene bank, as described above, revealed no significant homologies to 25 of these clones, hereinafter referred to as P5, P8, P9, P18, P20, P30, P34, P36, P38, P39, P42, P49, P50, P53, P55, P60, P64, P65, P73, P75, P76, P79 and P84. The determined cDNA sequences for these clones are provided in SEQ ID NO: 41-45, 47-52 and 54-65, respectively. P29, P47, P68, P80 and P82 (SEQ ID NO: 46, 53 and 66-68, respectively) were found to show some degree of homology to previously identified DNA sequences. To the best of the inventors' knowledge, none of these sequences have been previously shown to be present in prostate.

Further studies using the PCR-based methodology described above resulted in the isolation of more than 180 additional clones, of which 23 clones were found to show no significant homologies to known sequences. The determined cDNA sequences for these clones are provided in SEQ ID NO: 115-123, 127, 131, 137, 145, 147-151, 153, 156-158 and 160. Twenty-three clones (SEQ ID NO: 124-126, 128-130, 132-136, 138-144, 146, 152, 154, 155 and 159) were found to show some homology to previously identified ESTs. An additional ten clones (SEQ ID NO: 161-170) were found to have some degree of homology to known genes. Larger cDNA clones containing the P20 sequence represent splice variants of a gene referred to as P703P. The determined DNA sequence for the variants referred to as DE1, DE13 and DE14 are provided in SEQ ID NOS: 171, 175 and 177, respectively, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 172, 176 and 178, respectively. The determined cDNA sequence for an extended spliced form of P703 is provided in SEQ ID NO: 225. The DNA sequences for the splice variants referred to as DE2 and DE6 are provided in SEQ ID NOS: 173 and 174, respectively.

mRNA Expression levels for representative clones in tumor tissues (prostate (n=5), breast (n=2), colon and lung) normal tissues (prostate (n=5), colon, kidney, liver, lung (n=2), ovary (n=2), skeletal muscle, skin, stomach, small intestine and brain), and activated and non-activated PBMC was determined by RT-PCR as described above. Expression was examined in one sample of each tissue type unless otherwise indicated.

P9 was found to be highly expressed in normal prostate and prostate tumor compared to all normal tissues tested except for normal colon which showed comparable expression. P20, a portion of the P703P gene, was found to be highly expressed in normal prostate and prostate tumor, compared to all twelve normal tissues tested. A modest increase in expression of P20 in breast tumor (n=2), colon tumor and lung tumor was seen compared to all normal tissues except lung (1 of

2). Increased expression of P18 was found in normal prostate, prostate tumor and breast tumor compared to other normal tissues except lung and stomach. A modest increase in expression of P5 was observed in normal prostate compared to most other normal tissues. However, some elevated expression was seen in normal lung and PBMC. Elevated expression of P5 was also observed in prostate tumors (2 of 5), breast tumor and one lung tumor sample. For P30, similar expression levels were seen in normal prostate and prostate tumor, compared to six of twelve other normal tissues tested. Increased expression was seen in breast tumors, one lung tumor sample and one colon tumor sample, and also in normal PBMC. P29 was found to be over-expressed in prostate tumor (5 of 5) and normal prostate (5 of 5) compared to the majority of normal tissues. However, substantial expression of P29 was observed in normal colon and normal lung (2 of 2). P80 was found to be over-expressed in prostate tumor (5 of 5) and normal prostate (5 of 5) compared to all other normal tissues tested, with increased expression also being seen in colon tumor.

Further studies resulted in the isolation of twelve additional clones, hereinafter referred to as 10-d8, 10-h10, 11-c8, 7-g6, 8-b5, 8-b6, 8-d4, 8-d9, 8-g3, 8-h11, 9-f12 and 9-f3. The determined DNA sequences for 10-d8, 10-h10, 11-c8, 8-d4, 8-d9, 8-h11, 9-f12 and 9-f3 are provided in SEQ ID NO: 207, 208, 209, 216, 217, 220, 221 and 222, respectively. The determined forward and reverse DNA sequences for 7-g6, 8-b5, 8-b6 and 8-g3 are provided in SEQ ID NO: 210 and 211; 212 and 213; 214 and 215; and 218 and 219, respectively. Comparison of these sequences with those in the gene bank revealed no significant homologies to the sequence of 9-f3. The clones 10-d8, 11-c8 and 8-h11 were found to show some homology to previously isolated ESTs, while 10-h10, 8-b5, 8-b6, 8-d4, 8-d9, 8-g3 and 9-f12 were found to show some homology to previously identified genes. Further characterization of 7-G6 and 8-G3 showed identity to the known genes PAP and PSA, respectively.

mRNA expression levels for these clones were determined using the micro-array technology described above. The clones 7-G6, 8-G3, 8-B5, 8-B6, 8-D4, 8-D9, 9-F3, 9-F12, 9-H3, 10-A2, 10-A4, 11-C9 and 11-F2 were found to be over-expressed in prostate tumor and normal prostate, with expression in other tissues tested being low or undetectable. Increased expression of 8-F11 was seen in prostate tumor and normal prostate, bladder, skeletal muscle and colon. Increased expression of 10-H10 was seen in prostate tumor and normal prostate, bladder, lung, colon, brain and large intestine. Increased expression of 9-B1 was seen in prostate tumor, breast tumor, and normal prostate, salivary gland, large intestine and skin, with increased expression of 11-C8 being seen in prostate tumor, and normal prostate and large intestine.

An additional cDNA fragment derived from the PCR-based normal prostate subtraction, described above, was found to be prostate specific by both micro-array technology and RT-PCR. The determined cDNA sequence of this clone (referred to as 9-A11) is provided in SEQ ID NO: 226. Comparison of this sequence with those in the public databases revealed 99% identity to the known gene HOXB13.

Further studies led to the isolation of the clones 8-C6 and 8-H7. The determined cDNA sequences for these clones are provided in SEQ ID NO: 227 and 228, respectively. These sequences were found to show some homology to previously isolated ESTs.

PCR and hybridization-based methodologies were employed to obtain longer cDNA sequences for clone P20 (also referred to as P703P), yielding three additional cDNA fragments that progressively extend the 5' end of the gene. These fragments, referred to as P703PDE5, P703P6.26, and P703PX-23 (SEQ ID NO: 326, 328 and 330, with the predicted corresponding amino acid sequences being provided in SEQ ID NO: 327, 329 and 331, respectively) contain additional 5' sequence. P703PDE5 was recovered by screening of a cDNA library (#141-26) with a portion of P703P as a probe. P703P6.26 was recovered from a mixture of three prostate tumor cDNAs and P703PX_23 was recovered from cDNA library (#438-48). Together, the additional sequences include all of the putative mature serine protease along with part of the putative signal sequence. The putative full-length cDNA sequence for P703P is provided in SEQ ID NO: 524, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 525.

Further studies using a PCR-based subtraction library of a prostate tumor pool subtracted against a pool of normal tissues (referred to as JP: PCR subtraction) resulted in the isolation of thirteen additional clones, seven of which did not share any significant homology to known GenBank sequences. The determined cDNA sequences for these seven clones (P711P, P712P, novel 23, P774P, P775P, P710P and P768P) are provided in SEQ ID NO: 307-311, 313 and 315, respectively. The remaining six clones (SEQ ID NO: 316 and 321-325) were shown to share some homology to known genes. By microarray analysis, all thirteen clones showed three or more fold over-expression in prostate tissues, including prostate tumors, BPH and normal prostate as compared to normal non-prostate tissues. Clones P711P, P712P, novel 23 and P768P showed over-expression in most prostate tumors and BPH tissues tested (n=29), and in the majority of normal prostate tissues (n=4), but background to low expression levels in all normal tissues. Clones P774P, P775P and P710P showed comparatively lower expression and expression in fewer prostate tumors and BPH samples, with negative to low expression in normal prostate.

The full-length cDNA for P711P was obtained by employing the partial sequence of SEQ ID NO: 307 to screen a prostate cDNA library. Specifically, a directionally cloned prostate cDNA library was prepared using standard techniques. One million colonies of this library were plated onto LB/Amp plates. Nylon membrane filters were used to lift these colonies, and the cDNAs which were picked up by these filters were denatured and cross-linked to the filters by UV light. The P711P cDNA fragment of SEQ ID NO: 307 was radio-labeled and used to hybridize with these filters. Positive clones were selected, and cDNAs were prepared and sequenced using an automatic Perkin Elmer/Applied Biosystems sequencer. The determined full-length sequence of P711P is provided in SEQ ID NO: 382, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 383.

Using PCR and hybridization-based methodologies, additional cDNA sequence information was derived for two clones described above, 11-C9 and 9-F3, herein after referred to as P707P and P714P, respectively (SEQ ID NO: 333 and 334). After comparison with the most recent GenBank, P707P was found to be a splice variant of the known gene HoxB13. In contrast, no significant homologies to P714P were found.

Clones 8-B3, P89, P98, P130 and P201 (as disclosed in U.S. Patent Application No. 09/020,956, filed February 9, 1998) were found to be contained within one contiguous sequence, referred to as P705P (SEQ ID NO: 335, with the predicted amino acid sequence provided in SEQ ID NO: 336), which was determined to be a splice variant of the known gene NKX 3.1.

Further studies on P775P resulted in the isolation of four additional sequences (SEQ ID NO: 473-476) which are all splice variants of the P775P gene. The sequence of SEQ ID NO: 474 was found to contain two open reading frames (ORFs). The predicted amino acid sequences encoded by these ORFs are provided in SEQ ID NO: 477 and 478. The cDNA sequence of SEQ ID NO: 475 was found to contain an ORF which encodes the amino acid sequence of SEQ ID NO: 479. The cDNA sequence of SEQ ID NO: 473 was found to contain four ORFs. The predicted amino acid sequences encoded by these ORFs are provided in SEQ ID NO: 480-483.

Subsequent studies led to the identification of a genomic region on chromosome 22q11.2, known as the Cat Eye Syndrome region, that contains the five prostate genes P704P, P712P, P774P, P775P and B305D. The relative location of each of these five genes within the genomic region is shown in Fig. 10. This region may therefore be associated with malignant tumors, and other potential tumor genes may be contained within this region. These studies also led

to the identification of a potential open reading frame (ORF) for P775P (provided in SEQ ID NO: 533), which encodes the amino acid sequence of SEQ ID NO: 534.

EXAMPLE 4

SYNTHESIS OF POLYPEPTIDES

Polypeptides may be synthesized on a Perkin Elmer/Applied Biosystems 430A peptide synthesizer using FMOC chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray or other types of mass spectrometry and by amino acid analysis.

EXAMPLE 5

FURTHER ISOLATION AND CHARACTERIZATION OF PROSTATE-SPECIFIC POLYPEPTIDES BY PCR-BASED SUBTRACTION

A cDNA library generated from prostate primary tumor mRNA as described above was subtracted with cDNA from normal prostate. The subtraction was performed using a PCR-based protocol (Clontech), which was modified to generate larger fragments. Within this protocol, tester and driver double stranded cDNA were separately digested with five restriction enzymes that recognize six-nucleotide restriction sites (MluI, MscI, PvuII, SalI and StuI). This digestion resulted in an average cDNA size of 600 bp, rather than the average size of 300 bp that results from digestion with RsaI according to the Clontech protocol. This modification did not affect the

subtraction efficiency. Two tester populations were then created with different adapters, and the driver library remained without adapters.

The tester and driver libraries were then hybridized using excess driver cDNA. In the first hybridization step, driver was separately hybridized with each of the two tester cDNA populations. This resulted in populations of (a) unhybridized tester cDNAs, (b) tester cDNAs hybridized to other tester cDNAs, (c) tester cDNAs hybridized to driver cDNAs and (d) unhybridized driver cDNAs. The two separate hybridization reactions were then combined, and rehybridized in the presence of additional denatured driver cDNA. Following this second hybridization, in addition to populations (a) through (d), a fifth population (e) was generated in which tester cDNA with one adapter hybridized to tester cDNA with the second adapter. Accordingly, the second hybridization step resulted in enrichment of differentially expressed sequences which could be used as templates for PCR amplification with adaptor-specific primers.

The ends were then filled in, and PCR amplification was performed using adaptor-specific primers. Only population (e), which contained tester cDNA that did not hybridize to driver cDNA, was amplified exponentially. A second PCR amplification step was then performed, to reduce background and further enrich differentially expressed sequences.

This PCR-based subtraction technique normalizes differentially expressed cDNAs so that rare transcripts that are overexpressed in prostate tumor tissue may be recoverable. Such transcripts would be difficult to recover by traditional subtraction methods.

In addition to genes known to be overexpressed in prostate tumor, seventy-seven further clones were identified. Sequences of these partial cDNAs are provided in SEQ ID NO: 29 to 305. Most of these clones had no significant homology to database sequences. Exceptions were JPTPN23 (SEQ ID NO: 231; similarity to pig valosin-containing protein), JPTPN30 (SEQ ID NO: 234; similarity to rat mRNA for proteasome subunit), JPTPN45 (SEQ ID NO: 243; similarity to rat *norvegicus* cytosolic NADP-dependent isocitrate dehydrogenase), JPTPN46 (SEQ ID NO: 244; similarity to human subclone H8 4 d4 DNA sequence), JP1D6 (SEQ ID NO: 265; similarity to *G. gallus* dynein light chain-A), JP8D6 (SEQ ID NO: 288; similarity to human BAC clone RG016J04), JP8F5 (SEQ ID NO: 289; similarity to human subclone H8 3 b5 DNA sequence), and JP8E9 (SEQ ID NO: 299; similarity to human Alu sequence).

Additional studies using the PCR-based subtraction library consisting of a prostate tumor pool subtracted against a normal prostate pool (referred to as PT-PN PCR subtraction) yielded three additional clones. Comparison of the cDNA sequences of these clones with the most

recent release of GenBank revealed no significant homologies to the two clones referred to as P715P and P767P (SEQ ID NO: 312 and 314). The remaining clone was found to show some homology to the known gene KIAA0056 (SEQ ID NO: 318). Using microarray analysis to measure mRNA expression levels in various tissues, all three clones were found to be over-expressed in prostate tumors and BPH tissues. Specifically, clone P715P was over-expressed in most prostate tumors and BPH tissues by a factor of three or greater, with elevated expression seen in the majority of normal prostate samples and in fetal tissue, but negative to low expression in all other normal tissues. Clone P767P was over-expressed in several prostate tumors and BPH tissues, with moderate expression levels in half of the normal prostate samples, and background to low expression in all other normal tissues tested.

Further analysis, by microarray as described above, of the PT-PN PCR subtraction library and of a DNA subtraction library containing cDNA from prostate tumor subtracted with a pool of normal tissue cDNAs, led to the isolation of 27 additional clones (SEQ ID NO: 340-365 and 381) which were determined to be over-expressed in prostate tumor. The clones of SEQ ID NO: 341, 342, 345, 347, 348, 349, 351, 355-359, 361, 362 and 364 were also found to be expressed in normal prostate. Expression of all 26 clones in a variety of normal tissues was found to be low or undetectable, with the exception of P544S (SEQ ID NO: 356) which was found to be expressed in small intestine. Of the 26 clones, 10 (SEQ ID NO: 340-349) were found to show some homology to previously identified sequences. No significant homologies were found to the clones of SEQ ID NO: 350, 351 and 353-365.

Further studies on the clone of SEQ ID NO: 352 (referred to as P790P) led to the isolation of the full-length cDNA sequence of SEQ ID NO: 526. The corresponding predicted amino acid is provided in SEQ ID NO: 527. Data from two quantitative PCR experiments indicated that P790P is over-expressed in 11/15 tested prostate tumor samples and is expressed at low levels in spinal cord, with no expression being seen in all other normal samples tested. Data from further PCR experiments and microarray experiments showed over-expression in normal prostate and prostate tumor with little or no expression in other tissues tested. P790P was subsequently found to show significant homology to a previously identified G-protein coupled prostate tissue receptor.

EXAMPLE 6

PEPTIDE PRIMING OF MICE AND PROPAGATION OF CTL LINES

6.1. This Example illustrates the preparation of a CTL cell line specific for cells expressing the P502S gene.

Mice expressing the transgene for human HLA A2Kb (provided by Dr L. Sherman, The Scripps Research Institute, La Jolla, CA) were immunized with P2S#12 peptide (VLGWVAEL; SEQ ID NO: 306), which is derived from the P502S gene (also referred to herein as J1-17, SEQ ID NO: 8), as described by Theobald et al., *Proc. Natl. Acad. Sci. USA 92*:11993-11997, 1995 with the following modifications. Mice were immunized with 100µg of P2S#12 and 120µg of an I-A^b binding peptide derived from hepatitis B Virus protein emulsified in incomplete Freund's adjuvant. Three weeks later these mice were sacrificed and using a nylon mesh single cell suspensions prepared. Cells were then resuspended at 6 x 10^6 cells/ml in complete media (RPMI-1640; Gibco BRL, Gaithersburg, MD) containing 10% FCS, 2mM Glutamine (Gibco BRL), sodium pyruvate (Gibco BRL), non-essential amino acids (Gibco BRL), 2 x 10^{-5} M 2-mercaptoethanol, 50U/ml penicillin and streptomycin, and cultured in the presence of irradiated (3000 rads) P2S#12-pulsed (5mg/ml P2S#12 and 10mg/ml β2-microglobulin) LPS blasts (A2 transgenic spleens cells cultured in the presence of 7µg/ml dextran sulfate and 25µg/ml LPS for 3 days). Six days later, cells (5 x 10^5/ml) were restimulated with 2.5 x 10^6/ml peptide pulsed irradiated (20,000 rads) EL4A2Kb cells (Sherman et al, *Science 258*:815-818, 1992) and 3 x 10^6/ml A2 transgenic spleen feeder cells. Cells were cultured in the presence of 20U/ml IL-2. Cells continued to be restimulated on a weekly basis as described, in preparation for cloning the line.

P2S#12 line was cloned by limiting dilution analysis with peptide pulsed EL4 A2Kb tumor cells (1 x 10^4 cells/ well) as stimulators and A2 transgenic spleen cells as feeders (5 x 10^5 cells/ well) grown in the presence of 30U/ml IL-2. On day 14, cells were restimulated as before. On day 21, clones that were growing were isolated and maintained in culture. Several of these clones demonstrated significantly higher reactivity (lysis) against human fibroblasts (HLA A2Kb expressing) transduced with P502S than against control fibroblasts. An example is presented in Figure 1.

This data indicates that P2S #12 represents a naturally processed epitope of the P502S protein that is expressed in the context of the human HLA A2Kb molecule.

6.2. This Example illustrates the preparation of murine CTL lines and CTL clones specific for cells expressing the P501S gene.

This series of experiments were performed similarly to that described above. Mice were immunized with the P1S#10 peptide (SEQ ID NO: 337), which is derived from the P501S gene (also referred to herein as L1-12, SEQ ID NO: 110). The P1S#10 peptide was derived by analysis of the predicted polypeptide sequence for P501S for potential HLA-A2 binding sequences as defined by published HLA-A2 binding motifs (Parker, KC, *et al, J. Immunol.*, 152:163, 1994). P1S#10 peptide was synthesized as described in Example 4, and empirically tested for HLA-A2 binding using a T cell based competition assay. Predicted A2 binding peptides were tested for their ability to compete HLA-A2 specific peptide presentation to an HLA-A2 restricted CTL clone (D150M58), which is specific for the HLA-A2 binding influenza matrix peptide fluM58. D150M58 CTL secretes TNF in response to self-presentation of peptide fluM58. In the competition assay, test peptides at 100-200 μg/ml were added to cultures of D150M58 CTL in order to bind HLA-A2 on the CTL. After thirty minutes, CTL cultured with test peptides, or control peptides, were tested for their antigen dose response to the fluM58 peptide in a standard TNF bioassay. As shown in Figure 3, peptide P1S#10 competes HLA-A2 restricted presentation of fluM58, demonstrating that peptide P1S#10 binds HLA-A2.

Mice expressing the transgene for human HLA A2Kb were immunized as described by Theobald et al. (*Proc. Natl. Acad. Sci. USA 92*:11993-11997, 1995) with the following modifications. Mice were immunized with 62.5μg of P1S #10 and 120μg of an I-A^b binding peptide derived from Hepatitis B Virus protein emulsified in incomplete Freund's adjuvant. Three weeks later these mice were sacrificed and single cell suspensions prepared using a nylon mesh. Cells were then resuspended at 6 x 10^6 cells/ml in complete media (as described above) and cultured in the presence of irradiated (3000 rads) P1S#10-pulsed (2μg/ml P1S#10 and 10mg/ml β2-microglobulin) LPS blasts (A2 transgenic spleens cells cultured in the presence of 7μg/ml dextran sulfate and 25μg/ml LPS for 3 days). Six days later cells (5 x 10^5/ml) were restimulated with 2.5 x 10^6/ml peptide-pulsed irradiated (20,000 rads) EL4A2Kb cells, as described above, and 3 x 10^6/ml A2 transgenic spleen feeder cells. Cells were cultured in the presence of 20 U/ml IL-2. Cells were restimulated on a weekly basis in preparation for cloning. After three rounds of *in vitro* stimulations, one line was generated that recognized P1S#10-pulsed Jurkat A2Kb targets and P501S-transduced Jurkat targets as shown in Figure 4.

A P1S#10-specific CTL line was cloned by limiting dilution analysis with peptide pulsed EL4 A2Kb tumor cells (1 x 10^4 cells/ well) as stimulators and A2 transgenic spleen cells as feeders (5 x 10^5 cells/ well) grown in the presence of 30U/ml IL-2. On day 14, cells were restimulated as before. On day 21, viable clones were isolated and maintained in culture. As shown in Figure 5, five of these clones demonstrated specific cytolytic reactivity against P501S-transduced Jurkat A2Kb targets. This data indicates that P1S#10 represents a naturally processed epitope of the P501S protein that is expressed in the context of the human HLA-A2.1 molecule.

EXAMPLE 7

PRIMING OF CTL *IN VIVO* USING NAKED DNA IMMUNIZATION WITH A PROSTATE ANTIGEN

The prostate-specific antigen L1-12, as described above, is also referred to as P501S. HLA A2Kb Tg mice (provided by Dr L. Sherman, The Scripps Research Institute, La Jolla, CA) were immunized with 100 μg P501S in the vector VR1012 either intramuscularly or intradermally. The mice were immunized three times, with a two week interval between immunizations. Two weeks after the last immunization, immune spleen cells were cultured with Jurkat A2Kb-P501S transduced stimulator cells. CTL lines were stimulated weekly. After two weeks of *in vitro* stimulation, CTL activity was assessed against P501S transduced targets. Two out of 8 mice developed strong anti-P501S CTL responses. These results demonstrate that P501S contains at least one naturally processed HLA-A2-restricted CTL epitope.

EXAMPLE 8

ABILITY OF HUMAN T CELLS TO RECOGNIZE PROSTATE-SPECIFIC POLYPEPTIDES

This Example illustrates the ability of T cells specific for a prostate tumor polypeptide to recognize human tumor.

Human $CD8^+$ T cells were primed *in vitro* to the P2S-12 peptide (SEQ ID NO: 306) derived from P502S (also referred to as J1-17) using dendritic cells according to the protocol of Van Tsai et al. (*Critical Reviews in Immunology 18*:65-75, 1998). The resulting $CD8^+$ T cell microcultures were tested for their ability to recognize the P2S-12 peptide presented by autologous fibroblasts or fibroblasts which were transduced to express the P502S gene in a γ-interferon

ELISPOT assay (*see* Lalvani et al., *J. Exp. Med. 186*:859-865, 1997). Briefly, titrating numbers of T cells were assayed in duplicate on 10^4 fibroblasts in the presence of 3 μg/ml human β_2-microglobulin and 1 μg/ml P2S-12 peptide or control E75 peptide. In addition, T cells were simultaneously assayed on autologous fibroblasts transduced with the P502S gene or as a control, fibroblasts transduced with HER-2/*neu*. Prior to the assay, the fibroblasts were treated with 10 ng/ml γ-interferon for 48 hours to upregulate class I MHC expression. One of the microcultures (#5) demonstrated strong recognition of both peptide pulsed fibroblasts as well as transduced fibroblasts in a γ-interferon ELISPOT assay. Figure 2A demonstrates that there was a strong increase in the number of γ-interferon spots with increasing numbers of T cells on fibroblasts pulsed with the P2S-12 peptide (solid bars) but not with the control E75 peptide (open bars). This shows the ability of these T cells to specifically recognize the P2S-12 peptide. As shown in Figure 2B, this microculture also demonstrated an increase in the number of γ-interferon spots with increasing numbers of T cells on fibroblasts transduced to express the P502S gene but not the HER-2/*neu* gene. These results provide additional confirmatory evidence that the P2S-12 peptide is a naturally processed epitope of the P502S protein. Furthermore, this also demonstrates that there exists in the human T cell repertoire, high affinity T cells which are capable of recognizing this epitope. These T cells should also be capable of recognizing human tumors which express the P502S gene.

EXAMPLE 9

ELICITATION OF PROSTATE ANTIGEN-SPECIFIC CTL RESPONSES IN HUMAN BLOOD

This Example illustrates the ability of a prostate-specific antigen to elicit a CTL response in blood of normal humans.

Autologous dendritic cells (DC) were differentiated from monocyte cultures derived from PBMC of normal donors by growth for five days in RPMI medium containing 10% human serum, 50 ng/ml GMCSF and 30 ng/ml IL-4. Following culture, DC were infected overnight with recombinant P501S-expressing vaccinia virus at an M.O.I. of 5 and matured for 8 hours by the addition of 2 micrograms/ml CD40 ligand. Virus was inactivated by UV irradiation, $CD8^+$ cells were isolated by positive selection using magnetic beads, and priming cultures were initiated in 24-well plates. Following five stimulation cycles using autologous fibroblasts retrovirally transduced

to express P501S and CD80, CD8+ lines were identified that specifically produced interferon-gamma when stimulated with autologous P501S-transduced fibroblasts. The P501S-specific activity of cell line 3A-1 could be maintained following additional stimulation cycles on autologous B-LCL transduced with P501S. Line 3A-1 was shown to specifically recognize autologous B-LCL transduced to express P501S, but not EGFP-transduced autologous B-LCL, as measured by cytotoxicity assays (^{51}Cr release) and interferon-gamma production (Interferon-gamma Elispot; *see* above and Lalvani et al., *J. Exp. Med. 186*:859-865, 1997). The results of these assays are presented in Figures 6A and 6B.

EXAMPLE 10

IDENTIFICATION OF A NATURALLY PROCESSED CTL EPITOPE CONTAINED WITHIN A PROSTATE-SPECIFIC ANTIGEN

The 9-mer peptide p5 (SEQ ID NO: 338) was derived from the P703P antigen (also referred to as P20). The p5 peptide is immunogenic in human HLA-A2 donors and is a naturally processed epitope. Antigen specific human CD8+ T cells can be primed following repeated *in vitro* stimulations with monocytes pulsed with p5 peptide. These CTL specifically recognize p5-pulsed and P703P-transduced target cells in both ELISPOT (as described above) and chromium release assays. Additionally, immunization of HLA-A2Kb transgenic mice with p5 leads to the generation of CTL lines which recognize a variety of HLA-A2Kb or HLA-A2 transduced target cells expressing P703P.

Initial studies demonstrating that p5 is a naturally processed epitope were done using HLA-A2Kb transgenic mice. HLA-A2Kb transgenic mice were immunized subcutaneously in the footpad with 100 μg of p5 peptide together with 140 μg of hepatitis B virus core peptide (a Th peptide) in Freund's incomplete adjuvant. Three weeks post immunization, spleen cells from immunized mice were stimulated *in vitro* with peptide-pulsed LPS blasts. CTL activity was assessed by chromium release assay five days after primary *in vitro* stimulation. Retrovirally transduced cells expressing the control antigen P703P and HLA-A2Kb were used as targets. CTL lines that specifically recognized both p5-pulsed targets as well as P703P-expressing targets were identified.

Human *in vitro* priming experiments demonstrated that the p5 peptide is immunogenic in humans. Dendritic cells (DC) were differentiated from monocyte cultures derived

from PBMC of normal human donors by culturing for five days in RPMI medium containing 10% human serum, 50 ng/ml human GM-CSF and 30 ng/ml human IL-4. Following culture, the DC were pulsed with 1 ug/ml p5 peptide and cultured with CD8+ T cell enriched PBMC. CTL lines were restimulated on a weekly basis with p5-pulsed monocytes. Five to six weeks after initiation of the CTL cultures, CTL recognition of p5-pulsed target cells was demonstrated. CTL were additionally shown to recognize human cells transduced to express P703P, demonstrating that p5 is a naturally processed epitope.

EXAMPLE 11

EXPRESSION OF A BREAST TUMOR-DERIVED ANTIGEN IN PROSTATE

Isolation of the antigen B305D from breast tumor by differential display is described in US Patent Application No. 08/700,014, filed August 20, 1996. Several different splice forms of this antigen were isolated. The determined cDNA sequences for these splice forms are provided in SEQ ID NO: 366-375, with the predicted amino acid sequences corresponding to the sequences of SEQ ID NO: 292, 298 and 301-303 being provided in SEQ ID NO: 299-306, respectively. In further studies, a splice variant of the cDNA sequence of SEQ ID NO: 366 was isolated which was found to contain an additional guanine residue at position 884 (SEQ ID NO: 530), leading to a frameshift in the open reading frame. The determined DNA sequence of this ORF is provided in SEQ ID NO: 531. This frameshift generates a protein sequence (provided in SEQ ID NO: 532) of 293 amino acids that contains the C-terminal domain common to the other isoforms of B305D but that differs in the N-terminal region.

The expression levels of B305D in a variety of tumor and normal tissues were examined by real time PCR and by Northern analysis. The results indicated that B305D is highly expressed in breast tumor, prostate tumor, normal prostate and normal testes, with expression being low or undetectable in all other tissues examined (colon tumor, lung tumor, ovary tumor, and normal bone marrow, colon, kidney, liver, lung, ovary, skin, small intestine, stomach).

EXAMPLE 12

GENERATION OF HUMAN CTL *IN VITRO* USING WHOLE GENE PRIMING AND STIMULATION TECHNIQUES WITH PROSTATE-SPECIFIC ANTIGEN

Using *in vitro* whole-gene priming with P501S-vaccinia infected DC (see, for example, Yee et al, *The Journal of Immunology*, 157(9):4079-86, 1996), human CTL lines were derived that specifically recognize autologous fibroblasts transduced with P501S (also known as L1-12), as determined by interferon-γ ELISPOT analysis as described above. Using a panel of HLA-mismatched B-LCL lines transduced with P501S, these CTL lines were shown to be likely restricted to HLAB class I allele. Specifically, dendritic cells (DC) were differentiated from monocyte cultures derived from PBMC of normal human donors by growing for five days in RPMI medium containing 10% human serum, 50 ng/ml human GM-CSF and 30 ng/ml human IL-4. Following culture, DC were infected overnight with recombinant P501S vaccinia virus at a multiplicity of infection (M.O.I) of five, and matured overnight by the addition of 3 μg/ml CD40 ligand. Virus was inactivated by UV irradiation. CD8+ T cells were isolated using a magnetic bead system, and priming cultures were initiated using standard culture techniques. Cultures were restimulated every 7-10 days using autologous primary fibroblasts retrovirally transduced with P501S and CD80. Following four stimulation cycles, CD8+ T cell lines were identified that specifically produced interferon-γ when stimulated with P501S and CD80-transduced autologous fibroblasts. A panel of HLA-mismatched B-LCL lines transduced with P501S were generated to define the restriction allele of the response. By measuring interferon-γ in an ELISPOT assay, the P501S specific response was shown to be likely restricted by HLA B alleles. These results demonstrate that a CD8+ CTL response to P501S can be elicited.

To identify the epitope(s) recognized, cDNA encoding P501S was fragmented by various restriction digests, and sub-cloned into the retroviral expression vector pBIB-KS. Retroviral supernatants were generated by transfection of the helper packaging line Phoenix-Ampho. Supernatants were then used to transduce Jurkat/A2Kb cells for CTL screening. CTL were screened in IFN-gamma ELISPOT assays against these A2Kb targets transduced with the "library" of P501S fragments. Initial positive fragments P501S/H3 and P501S/F2 were sequenced and found to encode amino acids 106-553 and amino acids 136-547, respectively, of SEQ ID NO: 113. A truncation of H3 was made to encode amino acid residues 106-351 of SEQ ID NO: 113, which was unable to stimulate the CTL, thus localizing the epitope to amino acid residues 351-547. Additional fragments encoding amino acids 1-472 (Fragment A) and amino acids 1-351 (Fragment B) were also constructed. Fragment A but not Fragment B stimulated the CTL thus localizing the epitope to amino acid residues 351-472. Overlapping 20-mer and 18-mer peptides representing this region were tested by pulsing Jurkat/A2Kb cells versus CTL in an IFN-gamma assay. Only peptides

P501S-369(20) and P501S-369(18) stimulated the CTL. Nine-mer and 10-mer peptides representing this region were synthesized and similarly tested. Peptide P501S-370 (SEQ ID NO: 539) was the minimal 9-mer giving a strong response. Peptide P501S-376 (SEQ ID NO: 540) also gave a weak response, suggesting that it might represent a cross-reactive epitope.

In subsequent studies, the ability of primary human B cells transduced with P501S to prime MHC class I-restricted, P501S-specific, autologous CD8 T cells was examined. Primary B cells were derived from PBMC of a homozygous HLA-A2 donor by culture in CD40 ligand and IL-4, transduced at high frequency with recombinant P501S in the vector pBIB, and selected with blastocidin-S. For *in vitro* priming, purified CD8+ T cells were cultured with autologous CD40 ligand + IL-4 derived, P501S-transduced B cells in a 96-well microculture format. These CTL microcultures were re-stimulated with P501S-transduced B cells and then assayed for specificity. Following this initial screen, microcultures with significant signal above background were cloned on autologous EBV-transformed B cells (BLCL), also transduced with P501S. Using IFN-gamma ELISPOT for detection, several of these CD8 T cell clones were found to be specific for P501S, as demonstrated by reactivity to BLCL/P501S but not BLCL transduced with control antigen. It was further demonstrated that the anti-P501S CD8 T cell specificity is HLA-A2-restricted. First, antibody blocking experiments with anti-HLA-A,B,C monoclonal antibody (W6.32), anti-HLA-B,C monoclonal antibody (B1.23.2) and a control monoclonal antibody showed that only the anti-HLA-A,B,C antibody blocked recognition of P501S-expressing autologous BLCL. Secondly, the anti-P501S CTL also recognized an HLA-A2 matched, heterologous BLCL transduced with P501S, but not the corresponding EGFP transduced control BLCL.

EXAMPLE 13

IDENTIFICATION OF PROSTATE-SPECIFIC ANTIGENS BY MICROARRAY ANALYSIS

This Example describes the isolation of certain prostate-specific polypeptides from a prostate tumor cDNA library.

A human prostate tumor cDNA expression library as described above was screened using microarray analysis to identify clones that display at least a three fold over-expression in prostate tumor and/or normal prostate tissue, as compared to non-prostate normal tissues (not including testis). 372 clones were identified, and 319 were successfully sequenced. Table I presents a summary of these clones, which are shown in SEQ ID NOs:385-400. Of these sequences

SEQ ID NOs:386, 389, 390 and 392 correspond to novel genes, and SEQ ID NOs: 393 and 396 correspond to previously identified sequences. The others (SEQ ID NOs:385, 387, 388, 391, 394, 395 and 397-400) correspond to known sequences, as shown in Table I.

Table I

Summary of Prostate Tumor Antigens

Known Genes	Previously Identified Genes	Novel Genes
T-cell gamma chain	P504S	23379 (SEQ ID NO:389)
Kallikrein	P1000C	23399 (SEQ ID NO:392)
Vector	P501S	23320 (SEQ ID NO:386)
CGI-82 protein mRNA (23319; SEQ ID NO:385)	P503S	23381 (SEQ ID NO:390)
PSA	P510S	
Ald. 6 Dehyd.	P784P	
L-iditol-2 dehydrogenase (23376; SEQ ID NO:388)	P502S	
Ets transcription factor PDEF (22672; SEQ ID NO:398)	P706P	
hTGR (22678; SEQ ID NO:399)	19142.2, bangur.seq (22621; SEQ ID NO:396)	
KIAA0295(22685; SEQ ID NO:400)	5566.1 Wang (23404; SEQ ID NO:393)	
Prostatic Acid Phosphatase(22655; SEQ ID NO:397)	P712P	
transglutaminase (22611; SEQ ID NO:395)	P778P	
HDLBP (23508; SEQ ID NO:394)		
CGI-69 Protein(23367; SEQ ID NO:387)		
KIAA0122(23383; SEQ ID NO:391)		
TEEG		

CGI-82 showed 4.06 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 43% of prostate tumors, 25% normal prostate, not detected in other normal tissues tested. L-iditol-2 dehydrogenase showed 4.94 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 90% of prostate tumors, 100% of normal prostate, and not detected in other normal tissues tested. Ets transcription factor PDEF showed 5.55 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 47% prostate tumors, 25% normal prostate and not detected in other normal tissues tested. hTGR1 showed 9.11 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 63% of prostate tumors and is not detected in normal tissues tested including normal prostate. KIAA0295 showed 5.59 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 47% of prostate tumors, low to undetectable in normal tissues tested including normal prostate tissues. Prostatic acid phosphatase showed 9.14 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 67% of prostate tumors, 50% of normal prostate, and not detected in other normal tissues tested. Transglutaminase showed 14.84 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 30% of prostate tumors, 50% of normal prostate, and is not detected in other normal tissues tested. High density lipoprotein binding protein (HDLBP) showed 28.06 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 97% of prostate tumors, 75% of normal prostate, and is undetectable in all other normal tissues tested. CGI-69 showed 3.56 fold over-expression in prostate tissues as compared to other normal tissues tested. It is a low abundant gene, detected in more than 90% of prostate tumors, and in 75% normal prostate tissues. The expression of this gene in normal tissues was very low. KIAA0122 showed 4.24 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 57% of prostate tumors, it was undetectable in all normal tissues tested including normal prostate tissues. 19142.2 bangur showed 23.25 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 97% of prostate tumors and 100% of normal prostate. It was undetectable in other normal tissues tested. 5566.1 Wang showed 3.31 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 97% of prostate tumors, 75% normal prostate and was also over-expressed in normal bone marrow, pancreas, and activated PBMC. Novel clone 23379 showed 4.86 fold over-expression in prostate tissues as compared to other normal tissues tested. It was detectable in 97%

of prostate tumors and 75% normal prostate and is undetectable in all other normal tissues tested. Novel clone 23399 showed 4.09 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 27% of prostate tumors and was undetectable in all normal tissues tested including normal prostate tissues. Novel clone 23320 showed 3.15 fold over-expression in prostate tissues as compared to other normal tissues tested. It was detectable in all prostate tumors and 50% of normal prostate tissues. It was also expressed in normal colon and trachea. Other normal tissues do not express this gene at high level.

EXAMPLE 14

IDENTIFICATION OF PROSTATE-SPECIFIC ANTIGENS BY ELECTRONIC SUBTRACTION

This Example describes the use of an electronic subtraction technique to identify prostate-specific antigens.

Potential prostate-specific genes present in the GenBank human EST database were identified by electronic subtraction (similar to that described by Vasmatizis et al., *Proc. Natl. Acad. Sci. USA 95*:300-304, 1998). The sequences of EST clones (43,482) derived from various prostate libraries were obtained from the GenBank public human EST database. Each prostate EST sequence was used as a query sequence in a BLASTN (National Center for Biotechnology Information) search against the human EST database. All matches considered identical (length of matching sequence >100 base pairs, density of identical matches over this region > 70%) were grouped (aligned) together in a cluster. Clusters containing more than 200 ESTs were discarded since they probably represented repetitive elements or highly expressed genes such as those for ribosomal proteins. If two or more clusters shared common ESTs, those clusters were grouped together into a "supercluster," resulting in 4,345 prostate superclusters.

Records for the 479 human cDNA libraries represented in the GenBank release were downloaded to create a database of these cDNA library records. These 479 cDNA libraries were grouped into three groups: Plus (normal prostate and prostate tumor libraries, and breast cell line libraries, in which expression was desired), Minus (libraries from other normal adult tissues, in which expression was not desirable), and Other (libraries from fetal tissue, infant tissue, tissues found only in women, non-prostate tumors and cell lines other than prostate cell lines, in which

expression was considered to be irrelevant). A summary of these library groups is presented in Table II.

Table II

Prostate cDNA Libraries and ESTs

Library	# of Libraries	# of ESTs
Plus	25	43,482
Normal	11	18,875
Tumor	11	21,769
Cell lines	3	2,838
Minus	166	
Other	287	

Each supercluster was analyzed in terms of the ESTs within the supercluster. The tissue source of each EST clone was noted and used to classify the superclusters into four groups: Type 1- EST clones found in the Plus group libraries only; no expression detected in Minus or Other group libraries; Type 2- EST clones derived from the Plus and Other group libraries only; no expression detected in the Minus group; Type 3- EST clones derived from the Plus, Minus and Other group libraries, but the number of ESTs derived from the Plus group is higher than in either the Minus or Other groups; and Type 4- EST clones derived from Plus, Minus and Other group libraries, but the number derived from the Plus group is higher than the number derived from the Minus group. This analysis identified 4,345 breast clusters (*see* Table III). From these clusters, 3,172 EST clones were ordered from Research Genetics, Inc., and were received as frozen glycerol stocks in 96-well plates.

Table III

Prostate Cluster Summary

Type	# of Superclusters	# of ESTs Ordered
1	688	677
2	2899	2484
3	85	11
4	673	0
Total	4345	3172

The EST clone inserts were PCR-amplified using amino-linked PCR primers for Synteni microarray analysis. When more than one PCR product was obtained for a particular clone, that PCR product was not used for expression analysis. In total, 2,528 clones from the electronic subtraction method were analyzed by microarray analysis to identify electronic subtraction breast clones that had high levels of tumor vs. normal tissue mRNA. Such screens were performed using a Synteni (Palo Alto, CA) microarray, according to the manufacturer's instructions (and essentially as described by Schena et al., *Proc. Natl. Acad. Sci. USA 93*:10614-10619, 1996 and Heller et al., *Proc. Natl. Acad. Sci. USA 94*:2150-2155, 1997). Within these analyses, the clones were arrayed on the chip, which was then probed with fluorescent probes generated from normal and tumor prostate cDNA, as well as various other normal tissues. The slides were scanned and the fluorescence intensity was measured.

Clones with an expression ratio greater than 3 (*i.e.*, the level in prostate tumor and normal prostate mRNA was at least three times the level in other normal tissue mRNA) were identified as prostate tumor-specific sequences (Table IV). The sequences of these clones are provided in SEQ ID NO: 401-453, with certain novel sequences shown in SEQ ID NO: 407, 413, 416-419, 422, 426, 427 and 450.

Table IV

Prostate-tumor Specific Clones

SEQ ID NO.	Sequence Designation	Comments
401	22545	previously identified P1000C
402	22547	previously identified P704P
403	22548	known
404	22550	known
405	22551	PSA
406	22552	prostate secretory protein 94
407	22553	novel
408	22558	previously identified P509S
409	22562	glandular kallikrein
410	22565	previously identified P1000C
411	22567	PAP
412	22568	B1006C (breast tumor antigen)
413	22570	novel
414	22571	PSA
415	22572	previously identified P706P
416	22573	novel
417	22574	novel
418	22575	novel
419	22580	novel
420	22581	PAP
421	22582	prostatic secretory protein 94
422	22583	novel
423	22584	prostatic secretory protein 94
424	22585	prostatic secretory protein 94
425	22586	known
426	22587	novel
427	22588	novel
428	22589	PAP
429	22590	known
430	22591	PSA
431	22592	known
432	22593	Previously identified P777P
433	22594	T cell receptor gamma chain
434	22595	Previously identified P705P
435	22596	Previously identified P707P
436	22847	PAP
437	22848	known
438	22849	prostatic secretory protein 57
439	22851	PAP

440	22852	PAP
441	22853	PAP
442	22854	previously identified P509S
443	22855	previously identified P705P
444	22856	previously identified P774P
445	22857	PSA
446	23601	previously identified P777P
447	23602	PSA
448	23605	PSA
449	23606	PSA
450	23612	novel
451	23614	PSA
452	23618	previously identified P1000C
453	23622	previously identified P705P

EXAMPLE 15

FURTHER IDENTIFICATION OF PROSTATE-SPECIFIC ANTIGENS BY MICROARRAY ANALYSIS

This Example describes the isolation of additional prostate-specific polypeptides from a prostate tumor cDNA library.

A human prostate tumor cDNA expression library as described above was screened using microarray analysis to identify clones that display at least a three fold over-expression in prostate tumor and/or normal prostate tissue, as compared to non-prostate normal tissues (not including testis). 142 clones were identified and sequenced. Certain of these clones are shown in SEQ ID NO: 454-467. Of these sequences, SEQ ID NO: 459-461 represent novel genes. The others (SEQ ID NO: 454-458 and 461-467) correspond to known sequences.

EXAMPLE 16

FURTHER CHARACTERIZATION OF PROSTATE-SPECIFIC ANTIGEN P710P

This Example describes the full length cloning of P710P.

The prostate cDNA library described above was screened with the P710P fragment described above. One million colonies were plated on LB/Ampicillin plates. Nylon membrane

filters were used to lift these colonies, and the cDNAs picked up by these filters were then denatured and cross-linked to the filters by UV light. The P710P fragment was radiolabeled and used to hybridize with the filters. Positive cDNA clones were selected and their cDNAs recovered and sequenced by an automatic Perkin Elmer/Applied Biosystems Division Sequencer. Four sequences were obtained, and are presented in SEQ ID NO: 468-471 These sequences appear to represent different splice variants of the P710P gene.

EXAMPLE 17

PROTEIN EXPRESSION OF THE PROSTATE-SPECIFIC ANTIGEN P501S

This example describes the expression and purification of the prostate-specific antigen P501S in *E. coli*, baculovirus and mammalian cells.

a) Expression in *E. coli*

Expression of the full-length form of P501S was attempted by first cloning P501S without the leader sequence (amino acids 36-553 of SEQ ID NO: 113) downstream of the first 30 amino acids of the *M. tuberculosis* antigen Ra12 (SEQ ID NO: 484) in pET17b. Specifically, P501S DNA was used to perform PCR using the primers AW025 (SEQ ID NO: 485) and AW003 (SEQ ID NO: 486). AW025 is a sense cloning primer that contains a HindIII site. AW003 is an antisense cloning primer that contains an EcoRI site. DNA amplification was performed using 5 µl 10X Pfu buffer, 1 µl 20 mM dNTPs, 1 µl each of the PCR primers at 10 µM concentration, 40 µl water, 1 µl Pfu DNA polymerase (Stratagene, La Jolla, CA) and 1 µl DNA at 100 ng/µl. Denaturation at 95°C was performed for 30 sec, followed by 10 cycles of 95°C for 30 sec, 60°C for 1 min and by 72°C for 3 min. 20 cycles of 95°C for 30 sec, 65°C for 1 min and by 72°C for 3 min, and lastly by 1 cycle of 72°C for 10 min. The PCR product was cloned to Ra12m/pET17b using HindIII and EcoRI. The sequence of the resulting fusion construct (referred to as Ra12-P501S-F) was confirmed by DNA sequencing.

The fusion construct was transformed into BL21(DE3)pLysE, pLysS and CodonPlus *E. coli* (Stratagene) and grown overnight in LB broth with kanamycin. The resulting culture was induced with IPTG. Protein was transferred to PVDF membrane and blocked with 5% non-fat milk (in PBS-Tween buffer), washed three times and incubated with mouse anti-His tag antibody (Clontech) for 1 hour. The membrane was washed 3 times and probed with HRP-Protein A

(Zymed) for 30 min. Finally, the membrane was washed 3 times and developed with ECL (Amersham). No expression was detected by Western blot. Similarly, no expression was detected by Western blot when the Ra12-P501S-F fusion was used for expression in BL21CodonPlus by CE6 phage (Invitrogen).

An N-terminal fragment of P501S (amino acids 36-325 of SEQ ID NO: 113) was cloned down-stream of the first 30 amino acids of the *M. tuberculosis* antigen Ra12 in pET17b as follows. P501S DNA was used to perform PCR using the primers AW025 (SEQ ID NO: 485) and AW027 (SEQ ID NO: 487). AW027 is an antisense cloning primer that contains an EcoRI site and a stop codon. DNA amplification was performed essentially as described above. The resulting PCR product was cloned to Ra12 in pET17b at the HindIII and EcoRI sites. The fusion construct (referred to as Ra12-P501S-N) was confirmed by DNA sequencing.

The Ra12-P501S-N fusion construct was used for expression in BL21(DE3)pLysE, pLysS and CodonPlus, essentially as described above. Using Western blot analysis, protein bands were observed at the expected molecular weight of 36 kDa. Some high molecular weight bands were also observed, probably due to aggregation of the recombinant protein. No expression was detected by Western blot when the Ra12-P501S-F fusion was used for expression in BL21CodonPlus by CE6 phage.

A fusion construct comprising a C-terminal portion of P501S (amino acids 257-553 of SEQ ID NO: 113) located down-stream of the first 30 amino acids of the *M. tuberculosis* antigen Ra12 (SEQ ID NO: 484) was prepared as follows. P501S DNA was used to perform PCR using the primers AW026 (SEQ ID NO: 488) and AW003 (SEQ ID NO: 486). AW026 is a sense cloning primer that contains a HindIII site. DNA amplification was performed essentially as described above. The resulting PCR product was cloned to Ra12 in pET17b at the HindIII and EcoRI sites. The sequence for the fusion construct (referred to as Ra12-P501S-C) was confirmed.

The Ra12-P501S-C fusion construct was used for expression in BL21(DE3)pLysE, pLysS and CodonPlus, as described above. A small amount of protein was detected by Western blot, with some molecular weight aggregates also being observed. Expression was also detected by Western blot when the Ra12-P501S-C fusion was used for expression in BL21CodonPlus induced by CE6 phage.

b) Expression of P501S in Baculovirus

The Bac-to-Bac baculovirus expression system (BRL Life Technologies, Inc.) was used to express P501S protein in insect cells. Full-length P501S (SEQ ID NO: 113) was amplified by PCR and cloned into the XbaI site of the donor plasmid pFastBacI. The recombinant bacmid and baculovirus were prepared according to the manufacturer's isntructions. The recombinant baculovirus was amplified in Sf9 cells and the high titer viral stocks were utilized to infect High Five cells (Invitrogen) to make the recombinant protein. The identity of the full-length protein was confirmed by N-terminal sequencing of the recombinant protein and by Western blot analysis (Figure 7). Specifically, 0.6 million High Five cells in 6-well plates were infected with either the unrelated control virus BV/ECD_PD (lane 2), with recombinant baculovirus for P501S at different amounts or MOIs (lanes 4-8), or were uninfected (lane 3). Cell lysates were run on SDS-PAGE under reducing conditions and analyzed by Western blot with the anti-P501S monoclonal antibody P501S-10E3-G4D3 (prepared as described below). Lane 1 is the biotinylated protein molecular weight marker (BioLabs).

The localization of recombinant P501S in the insect cells was investigated as follows. The insect cells overexpressing P501S were fractionated into fractions of nucleus, mitochondria, membrane and cytosol. Equal amounts of protein from each fraction were analyzed by Western blot with a monoclonal antibody against P501S. Due to the scheme of fractionation, both nucleus and mitochondria fractions contain some plasma membrane components. However, the membrane fraction is basically free from mitochondria and nucleus. P501S was found to be present in all fractions that contain the membrane component, suggesting that P501S may be associated with plasma membrane of the insect cells expressing the recombinant protein.

c) Expression of P501S in mammalian cells

Full-length P501S (553AA) was cloned into various mammalian expression vectors, including pCEP4 (Invitrogen), pVR1012 (Vical, San Diego, CA) and a modified form of the retroviral vector pBMN, referred to as pBIB. Transfection of P501S/pCEP4 and P501S/pVR1012 into HEK293 fibroblasts was carried out using the Fugene transfection reagent (Boehringer Mannheim). Briefly, 2 ul of Fugene reagent was diluted into 100 ul of serum-free media and incubated at room temperature for 5-10 min. This mixture was added to 1 ug of P501S plasmid DNA, mixed briefly and incubated for 30 minutes at room temperature. The Fugene/DNA mixture

was added to cells and incubated for 24-48 hours. Expression of recombinant P501S in transfected HEK293 fibroblasts was detected by means of Western blot employing a monoclonal antibody to P501S.

Transfection of p501S/pCEP4 into CHO-K cells (American Type Culture Collection, Rockville, MD) was carried out using GenePorter transfection reagent (Gene Therapy Systems, San Diego, CA). Briefly, 15 µl of GenePorter was diluted in 500 µl of serum-free media and incubated at room temperature for 10 min. The GenePorter/media mixture was added to 2 µg of plasmid DNA that was diluted in 500 µl of serum-free media, mixed briefly and incubated for 30 min at room temperature. CHO-K cells were rinsed in PBS to remove serum proteins, and the GenePorter/DNA mix was added and incubated for 5 hours. The transfected cells were then fed an equal volume of 2x media and incubated for 24-48 hours.

FACS analysis of P501S transiently infected CHO-K cells, demonstrated surface expression of P501S. Expression was detected using rabbit polyclonal antisera raised against a P501S peptide, as described below. Flow cytometric analysis was performed using a FaCScan (Becton Dickinson), and the data were analyzed using the Cell Quest program.

EXAMPLE 18

PREPARATION AND CHARACTERIZATION OF ANTIBODIES AGAINST PROSTATE-SPECIFIC POLYPEPTIDES

a) Preparation and Characterization of Antibodies against P501S

A murine monoclonal antibody directed against the carboxy-terminus of the prostate-specific antigen P501S was prepared as follows.

A truncated fragment of P501S (amino acids 355-526 of SEQ ID NO: 113) was generated and cloned into the pET28b vector (Novagen) and expressed in *E. coli* as a thioredoxin fusion protein with a histidine tag. The trx-P501S fusion protein was purified by nickel chromatography, digested with thrombin to remove the trx fragment and further purified by an acid precipitation procedure followed by reverse phase HPLC.

Mice were immunized with truncated P501S protein. Serum bleeds from mice that potentially contained anti-P501S polyclonal sera were tested for P501S-specific reactivity using ELISA assays with purified P501S and trx-P501S proteins. Serum bleeds that appeared to react specifically with P501S were then screened for P501S reactivity by Western analysis. Mice that contained a P501S-specific antibody component were sacrificed and spleen cells were used to

generate anti-P501S antibody producing hybridomas using standard techniques. Hybridoma supernatants were tested for P501S-specific reactivity initially by ELISA, and subsequently by FACS analysis of reactivity with P501S transduced cells. Based on these results, a monoclonal hybridoma referred to as 10E3 was chosen for further subcloning. A number of subclones were generated, tested for specific reactivity to P501S using ELISA and typed for IgG isotype. The results of this analysis are shown below in Table V. Of the 16 subclones tested, the monoclonal antibody 10E3-G4-D3 was selected for further study.

Table V

Isotype analysis of murine anti-P501S monoclonal antibodies

Hybridoma clone	Isotype	Estimated [Ig] in supernatant (μg/ml)
4D11	IgG1	14.6
1G1	IgG1	0.6
4F6	IgG1	72
4H5	IgG1	13.8
4H5-E12	IgG1	10.7
4H5-EH2	IgG1	9.2
4H5-H2-A10	IgG1	10
4H5-H2-A3	IgG1	12.8
4H5-H2-A10-G6	IgG1	13.6
4H5-H2-B11	IgG1	12.3
10E3	IgG2a	3.4
10E3-D4	IgG2a	3.8
10E3-D4-G3	IgG2a	9.5
10E3-D4-G6	IgG2a	10.4
10E3-E7	IgG2a	6.5
8H12	IgG2a	0.6

The specificity of 10E3-G4-D3 for P501S was examined by FACS analysis. Specifically, cells were fixed (2% formaldehyde, 10 minutes), permeabilized (0.1% saponin, 10 minutes) and stained with 10E3-G4-D3 at 0.5 – 1 μg/ml, followed by incubation with a secondary, FITC-conjugated goat anti-mouse Ig antibody (Pharmingen, San Diego, CA). Cells were then analyzed for FITC fluorescence using an Excalibur fluorescence activated cell sorter. For FACS analysis of transduced cells, B-LCL were retrovirally transduced with P501S. For analysis of infected cells, B-LCL were infected with a vaccinia vector that expresses P501S. To demonstrate

specificity in these assays, B-LCL transduced with a different antigen (P703P) and uninfected B-LCL vectors were utilized. 10E3-G4-D3 was shown to bind with P501S-transduced B-LCL and also with P501S-infected B-LCL, but not with either uninfected cells or P703P-transduced cells.

To determine whether the epitope recognized by 10E3-G4-D3 was found on the surface or in an intracellular compartment of cells, B-LCL were transduced with P501S or HLA-B8 as a control antigen and either fixed and permeabilized as described above or directly stained with 10E3-G4-D3 and analyzed as above. Specific recognition of P501S by 10E3-G4-D3 was found to require permeabilization, suggesting that the epitope recognized by this antibody is intracellular.

The reactivity of 10E3-G4-D3 with the three prostate tumor cell lines Lncap, PC-3 and DU-145, which are known to express high, medium and very low levels of P501S, respectively, was examined by permeabilizing the cells and treating them as described above. Higher reactivity of 10E3-G4-D3 was seen with Lncap than with PC-3, which in turn showed higher reactivity that DU-145. These results are in agreement with the real time PCR and demonstrate that the antibody specifically recognizes P501S in these tumor cell lines and that the epitope recognized in prostate tumor cell lines is also intracellular.

Specificity of 10E3-G4-D3 for P501S was also demonstrated by Western blot analysis. Lysates from the prostate tumor cell lines Lncap, DU-145 and PC-3, from P501S-transiently transfected HEK293 cells, and from non-transfected HEK293 cells were generated. Western blot analysis of these lysates with 10E3-G4-D3 revealed a 46 kDa immunoreactive band in Lncap, PC-3 and P501S-transfected HEK cells, but not in DU-145 cells or non-transfected HEK293 cells. P501S mRNA expression is consistent with these results since semi-quantitative PCR analysis revealed that P501S mRNA is expressed in Lncap, to a lesser but detectable level in PC-3 and not at all in DU-145 cells. Bacterially expressed and purified recombinant P501S (referred to as P501SStr2) was recognized by 10E3-G4-D3 (24 kDa), as was full-length P501S that was transiently expressed in HEK293 cells using either the expression vector VR1012 or pCEP4. Although the predicted molecular weight of P501S is 60.5 kDa, both transfected and "native" P501S run at a slightly lower mobility due to its hydrophobic nature.

Immunohistochemical analysis was performed on prostate tumor and a panel of normal tissue sections (prostate, adrenal, breast, cervix, colon, duodenum, gall bladder, ileum, kidney, ovary, pancreas, parotid gland, skeletal muscle, spleen and testis). Tissue samples were fixed in formalin solution for 24 hours and embedded in paraffin before being sliced into 10 micron sections. Tissue sections were permeabilized and incubated with 10E3-G4-D3 antibody for 1 hr.

HRP-labeled anti-mouse followed by incubation with DAB chromogen was used to visualize P501S immunoreactivity. P501S was found to be highly expressed in both normal prostate and prostate tumor tissue but was not detected in any of the other tissues tested.

To identify the epitope recognized by 10E3-G4-D3, an epitope mapping approach was pursued. A series of 13 overlapping 20-21 mers (5 amino acid overlap; SEQ ID NO: 489-501) was synthesized that spanned the fragment of P501S used to generate 10E3-G4-D3. Flat bottom 96 well microtiter plates were coated with either the peptides or the P501S fragment used to immunize mice, at 1 microgram/ml for 2 hours at 37 °C. Wells were then aspirated and blocked with phosphate buffered saline containing 1% (w/v) BSA for 2 hours at room temperature, and subsequently washed in PBS containing 0.1% Tween 20 (PBST). Purified antibody 10E3-G4-D3 was added at 2 fold dilutions (1000 ng – 16 ng) in PBST and incubated for 30 minutes at room temperature. This was followed by washing 6 times with PBST and subsequently incubating with HRP-conjugated donkey anti-mouse IgG (H+L)Affinipure F(ab') fragment (Jackson Immunoresearch, West Grove, PA) at 1:20000 for 30 minutes. Plates were then washed and incubated for 15 minutes in tetramethyl benzidine. Reactions were stopped by the addition of 1N sulfuric acid and plates were read at 450 nm using an ELISA plate reader. As shown in Fig. 8, reactivity was seen with the peptide of SEQ ID NO: 496 (corresponding to amino acids 439-459 of P501S) and with the P501S fragment but not with the remaining peptides, demonstrating that the epitope recognized by 10E3-G4-D3 is localized to amino acids 439-459 of SEQ ID NO: 113.

In order to further evaluate the tissue specificity of P501S, multi-array immunohistochemical analysis was performed on approximately 4700 different human tissues encompassing all the major normal organs as well as neoplasias derived from these tissues. Sixty-five of these human tissue samples were of prostate origin. Tissue sections 0.6 mm in diameter were formalin-fixed and paraffin embedded. Samples were pretreated with HIER using 10 mM citrate buffer pH 6.0 and boiling for 10 min. Sections were stained with 10E3-G4-D3 and P501S immunoreactivity was visualized with HRP. All the 65 prostate tissues samples (5 normal, 55 untreated prostate tumors, 5 hormone refractory prostate tumors) were positive, showing distinct perinuclear staining. All other tissues examined were negative for P501S expression.

b) Preparation and Characterization of Antibodies against P503S

A fragment of P503S (amino acids 113-241 of SEQ ID NO: 114) was expressed and purified from bacteria essentially as described above for P501S and used to immunize both rabbits

and mice. Mouse monoclonal antibodies were isolated using standard hybridoma technology as described above. Rabbit monoclonal antibodies were isolated using Selected Lymphocyte Antibody Method (SLAM) technology at Immgenics Pharmaceuticals (Vancouver, BC, Canada). Table VI, below, lists the monoclonal antibodies that were developed against P503S.

Table VI

Antibody	Species
20D4	Rabbit
JA1	Rabbit
1A4	Mouse
1C3	Mouse
1C9	Mouse
1D12	Mouse
2A11	Mouse
2H9	Mouse
4H7	Mouse
8A8	Mouse
8D10	Mouse
9C12	Mouse
6D12	Mouse

The DNA sequences encoding the complementarity determining regions (CDRs) for the rabbit monoclonal antibodies 20D4 and JA1 were determined and are provided in SEQ ID NO: 502 and 503, respectively.

In order to better define the epitope binding region of each of the antibodies, a series of overlapping peptides were generated that span amino acids 109-213 of SEQ ID NO: 114. These peptides were used to epitope map the anti-P503S monoclonal antibodies by ELISA as follows. The recombinant fragment of P503S that was employed as the immunogen was used as a positive control. Ninety-six well microtiter plates were coated with either peptide or recombinant antigen at 20 ng/well overnight at 4 °C. Plates were aspirated and blocked with phosphate buffered saline containing 1% (w/v) BSA for 2 hours at room temperature then washed in PBS containing 0.1% Tween 20 (PBST). Purified rabbit monoclonal antibodies diluted in PBST were added to the wells and incubated for 30 min at room temperature. This was followed by washing 6 times with PBST and incubation with Protein-A HRP conjugate at a 1:2000 dilution for a further 30 min. Plates were washed six times in PBST and incubated with tetramethylbenzidine (TMB) substrate for a further

15 min. The reaction was stopped by the addition of 1N sulfuric acid and plates were read at 450 nm using at ELISA plate reader. ELISA with the mouse monoclonal antibodies was performed with supernatants from tissue culture run neat in the assay.

All of the antibodies bound to the recombinant P503S fragment, with the exception of the negative control SP2 supernatant. 20D4, JA1 and 1D12 bound strictly to peptide #2101 (SEQ ID NO: 504), which corresponds to amino acids 151-169 of SEQ ID NO: 114. 1C3 bound to peptide #2102 (SEQ ID NO: 505), which corresponds to amino acids 165-184 of SEQ ID NO: 114. 9C12 bound to peptide #2099 (SEQ ID NO: 522), which corresponds to amino acids 120-139 of SEQ ID NO: 114. The other antibodies bind to regions that were not examined in these studies.

Subsequent to epitope mapping, the antibodies were tested by FACS analysis on a cell line that stably expressed P503S to confirm that the antibodies bind to cell surface epitopes. Cells stably transfected with a control plasmid were employed as a negative control. Cells were stained live with no fixative. 0.5 ug of anti-P503S monoclonal antibody was added and cells were incubated on ice for 30 min before being washed twice and incubated with a FITC-labelled goat anti-rabbit or mouse secondary antibody for 20 min. After being washed twice, cells were analyzed with an Excalibur fluorescent activated cell sorter. The monoclonal antibodies 1C3, 1D12, 9C12, 20D4 and JA1, but not 8D3, were found to bind to a cell surface epitope of P503S.

In order to determine which tissues express P503S, immunohistochemical analysis was performed, essentially as described above, on a panel of normal tissues (prostate, adrenal, breast, cervix, colon, duodenum, gall bladder, ileum, kidney, ovary, pancreas, parotid gland, skeletal muscle, spleen and testis). HRP-labeled anti-mouse or anti-rabbit antibody followed by incubation with TMB was used to visualize P503S immunoreactivity. P503S was found to be highly expressed in prostate tissue, with lower levels of expression being observed in cervix, colon, ileum and kidney, and no expression being observed in adrenal, breast, duodenum, gall bladder, ovary, pancreas, parotid gland, skeletal muscle, spleen and testis.

Western blot analysis was used to characterize anti-P503S monoclonal antibody specificity. SDS-PAGE was performed on recombinant (rec) P503S expressed in and purified from bacteria and on lysates from HEK293 cells transfected with full length P503S. Protein was transferred to nitrocellulose and then Western blotted with each of the anti-P503S monoclonal antibodies (20D4, JA1, 1D12, 6D12 and 9C12) at an antibody concentration of 1 ug/ml. Protein was detected using horse radish peroxidase (HRP) conjugated to either a goat anti-mouse monoclonal antibody or to protein A-sepharose. The monoclonal antibody 20D4 detected the

appropriate molecular weight 14 kDa recombinant P503S (amino acids 113-241) and the 23.5 kDa species in the HEK293 cell lysates transfected with full length P503S. Other anti-P503S monoclonal antibodies displayed similar specificity by Western blot.

c) Preparation and Characterization of Antibodies against P703P

Rabbits were immunized with either a truncated (P703Ptr1; SEQ ID NO: 172) or full-length mature form (P703Pfl; SEQ ID NO: 523) of recombinant P703P protein was expressed in and purified from bacteria as described above. Affinity purified polyclonal antibody was generated using immunogen P703Pfl or P703Ptr1 attached to a solid support. Rabbit monoclonal antibodies were isolated using SLAM technology at Immgenics Pharmaceuticals. Table VII below lists both the polyclonal and monoclonal antibodies that were generated against P703P.

Table VII

Antibody	Immunogen	Species/type
Aff. Purif. P703P (truncated); #2594	P703Ptrl	Rabbit polyclonal
Aff. Purif. P703P (full length); #9245	P703Pfl	Rabbit polyclonal
2D4	P703Ptrl	Rabbit monoclonal
8H2	P703Ptrl	Rabbit monoclonal
7H8	P703Ptrl	Rabbit monoclonal

The DNA sequences encoding the complementarity determining regions (CDRs) for the rabbit monoclonal antibodies 8H2, 7H8 and 2D4 were determined and are provided in SEQ ID NO: 506-508, respectively.

Epitope mapping studies were performed as described above. Monoclonal antibodies 2D4 and 7H8 were found to specifically bind to the peptides of SEQ ID NO: 509 (corresponding to amino acids 145-159 of SEQ ID NO: 172) and SEQ ID NO: 510 (corresponding to amino acids 11-25 of SEQ ID NO: 172), respectively. The polyclonal antibody 2594 was found to bind to the peptides of SEQ ID NO: 511-514, with the polyclonal antibody 9427 binding to the peptides of SEQ ID NO: 515-517.

The specificity of the anti-P703P antibodies was determined by Western blot analysis as follows. SDS-PAGE was performed on (1) bacterially expressed recombinant antigen; (2) lysates of HEK293 cells and Ltk-/- cells either untransfected or transfected with a plasmid

expressing full length P703P; and (3) supernatant isolated from these cell cultures. Protein was transferred to nitrocellulose and then Western blotted using the anti-P703P polyclonal antibody #2594 at an antibody concentration of 1 ug/ml. Protein was detected using horse radish peroxidase (HRP) conjugated to an anti-rabbit antibody. A 35 kDa immunoreactive band could be observed with recombinant P703P. Recombinant P703P runs at a slightly higher molecular weight since it is epitope tagged. In lysates and supernatants from cells transfected with full length P703P, a 30 kDa band corresponding to P703P was observed. To assure specificity, lysates from HEK293 cells stably transfected with a control plasmid were also tested and were negative for P703P expression. Other anti-P703P antibodies showed similar results.

Immunohistochemical studies were performed as described above, using anti-P703P monoclonal antibody. P703P was found to be expressed at high levels in normal prostate and prostate tumor tissue but was not detectable in all other tissues tested (breast tumor, lung tumor and normal kidney).

EXAMPLE 19

CHARACTERIZATION OF CELL SURFACE EXPRESSION AND CHROMOSOME LOCALIZATION OF THE PROSTATE-SPECIFIC ANTIGEN P501S

This example describes studies demonstrating that the prostate-specific antigen P501S is expressed on the surface of cells, together with studies to determine the probable chromosomal location of P501S.

The protein P501S (SEQ ID NO: 113) is predicted to have 11 transmembrane domains. Based on the discovery that the epitope recognized by the anti-P501S monoclonal antibody 10E3-G4-D3 (described above in Example 17) is intracellular, it was predicted that following transmembrane determinants would allow the prediction of extracellular domains of P501S. Fig. 9 is a schematic representation of the P501S protein showing the predicted location of the transmembrane domains and the intracellular epitope described in Example 17. Underlined sequence represents the predicted transmembrane domains, bold sequence represents the predicted extracellular domains, and italized sequence represents the predicted intracellular domains. Sequence that is both bold and underlined represents sequence employed to generate polyclonal rabbit serum. The location of the transmembrane domains was predicted using HHMTOP as

described by Tusnady and Simon (Principles Governing Amino Acid Composition of Integral Membrane Proteins: Applications to Topology Prediction, *J. Mol. Biol. 283*:489-506, 1998).

Based on Fig. 9, the P501S domain flanked by the transmembrane domains corresponding to amino acids 274-295 and 323-342 is predicted to be extracellular. The peptide of SEQ ID NO: 518 corresponds to amino acids 306-320 of P501S and lies in the predicted extracellular domain. The peptide of SEQ ID NO: 519, which is identical to the peptide of SEQ ID NO: 518 with the exception of the substitution of the histidine with an asparginine, was synthesized as described above. A Cys-Gly was added to the C-terminus of the peptide to facilitate conjugation to the carrier protein. Cleavage of the peptide from the solid support was carried out using the following cleavage mixture: trifluoroacetic acid:ethanediol:thioanisol:water:phenol (40:1:2:2:3). After cleaving for two hours, the peptide was precipitated in cold ether. The peptide pellet was then dissolved in 10% v/v acetic acid and lyophilized prior to purification by C18 reverse phase hplc. A gradient of 5-60% acetonitrile (containing 0.05% TFA) in water (containing 0.05% TFA) was used to elute the peptide. The purity of the peptide was verified by hplc and mass spectrometry, and was determined to be >95%. The purified peptide was used to generate rabbit polyclonal antisera as described above.

Surface expression of P501S was examined by FACS analysis. Cells were stained with the polyclonal anti-P501S peptide serum at 10 μg/ml, washed, incubated with a secondary FITC-conjugated goat anti-rabbit Ig antibody (ICN), washed and analyzed for FITC fluorescence using an Excalibur fluorescence activated cell sorter. For FACS analysis of transduced cells, B-LCL were retrovirally transduced with P501S. To demonstrate specificity in these assays, B-LCL transduced with an irrelevant antigen (P703P) or nontransduced were stained in parallel. For FACS analysis of prostate tumor cell lines, Lncap, PC-3 and DU-145 were utilized. Prostate tumor cell lines were dissociated from tissue culture plates using cell dissociation medium and stained as above. All samples were treated with propidium iodide (PI) prior to FACS analysis, and data was obtained from PI-excluding (i.e. intact and non-permeabilized) cells. The rabbit polyclonal serum generated against the peptide of SEQ ID NO: 519 was shown to specifically recognize the surface of cells transduced to express P501S, demonstrating that the epitope recognized by the polyclonal serum is extracellular.

To determine biochemically if P501S is expressed on the cell surface, peripheral membranes from Lncap cells were isolated and subjected to Western blot analysis. Specifically, Lncap cells were lysed using a dounce homogenizer in 5 ml of homogenization buffer (250 mM

sucrose, 10 mM HEPES, 1mM EDTA, pH 8.0, 1 complete protease inhibitor tablet (Boehringer Mannheim)). Lysate samples were spun at 1000 g for 5 min at 4 °C. The supernatant was then spun at 8000g for 10 min at 4 °C. Supernatant from the 8000g spin was recovered and subjected to a 100,000g spin for 30 min at 4 °C to recover peripheral membrane. Samples were then separated by SDS-PAGE and Western blotted with the mouse monoclonal antibody 10E3-G4-D3 (described above in Example 17) using conditions described above. Recombinant purified P501S, as well as HEK293 cells transfected with and over-expressing P501S were included as positive controls for P501S detection. LCL cell lysate was included as a negative control. P501S could be detected in Lncap total cell lysate, the 8000g (internal membrane) fraction and also in the 100,000g (plasma membrane) fraction. These results indicate that P501S is expressed at, and localizes to, the peripheral membrane.

To demonstrate that the rabbit polyclonal antiserum generated to the peptide of SEQ ID NO: 519 specifically recognizes this peptide as well as the corresponding native peptide of SEQ ID NO: 518, ELISA analyses were performed. For these analyses, flat-bottomed 96 well microtiter plates were coated with either the peptide of SEQ ID NO: 519, the longer peptide of SEQ ID NO: 520 that spans the entire predicted extracellular domain, the peptide of SEQ ID NO: 521 which represents the epitope recognized by the P501S-specific antibody 10E3-G4-D3, or a P501S fragment (corresponding to amino acids 355-526 of SEQ ID NO: 113) that does not include the immunizing peptide sequence, at 1 µg/ml for 2 hours at 37 °C. Wells were aspirated, blocked with phosphate buffered saline containing 1% (w/v) BSA for 2 hours at room temperature and subsequently washed in PBS containing 0.1% Tween 20 (PBST). Purified anti-P501S polyclonal rabbit serum was added at 2 fold dilutions (1000 ng - 125 ng) in PBST and incubated for 30 min at room temperature. This was followed by washing 6 times with PBST and incubating with HRP-conjugated goat anti-rabbit IgG (H+L) Affinipure F(ab') fragment at 1:20000 for 30 min. Plates were then washed and incubated for 15 min in tetramethyl benzidine. Reactions were stopped by the addition of 1N sulfuric acid and plates were read at 450 nm using an ELISA plate reader. As shown in Fig. 11, the anti-P501S polyclonal rabbit serum specifically recognized the peptide of SEQ ID NO: 519 used in the immunization as well as the longer peptide of SEQ ID NO: 520, but did not recognize the irrelevant P501S-derived peptides and fragments.

In further studies, rabbits were immunized with peptides derived from the P501S sequence and predicted to be either extracellular or intracellular, as shown in Fig. 9. Polyclonal rabbit sera were isolated and polyclonal antibodies in the serum were purified, as described above.

To determine specific reactivity with P501S, FACS analysis was employed, utilizing either B-LCL transduced with P501S or the irrelevant antigen P703P, of B-LCL infected with vaccinia virus-expressing P501S. For surface expression, dead and non-intact cells were excluded from the analysis as described above. For intracellular staining, cells were fixed and permeabilized as described above. Rabbit polyclonal serum generated against the peptide of SEQ ID NO: 548, which corresponds to amino acids 181-198 of P501S, was found to recognize a surface epitope of P501S. Rabbit polyclonal serum generated against the peptide SEQ ID NO: 551, which corresponds to amino acids 543-553 of P501S, was found to recognize an epitope that was either potentially extracellular or intracellular since in different experiments intact or permeabilized cells were recognized by the polyclonal sera. Based on similar deductive reasoning, the sequences of SEQ ID NO: 541-547, 549 and 550, which correspond to amino acids 109-122, 539-553, 509-520, 37-54, 342-359, 295-323, 217-274, 143-160 and 75-88, respectively, of P501S, can be considered to be potential surface epitopes of P501S recognized by antibodies.

The chromosomal location of P501S was determined using the GeneBridge 4 Radiation Hybrid panel (Research Genetics). The PCR primers of SEQ ID NO: 528 and 529 were employed in PCR with DNA pools from the hybrid panel according to the manufacturer's directions. After 38 cycles of amplification, the reaction products were separated on a 1.2% agarose gel, and the results were analyzed through the Whitehead Institute/MIT Center for Genome Research web server (http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl) to determine the probable chromosomal location. Using this approach, P501S was mapped to the long arm of chromosome 1 at WI-9641 between q32 and q42. This region of chromosome 1 has been linked to prostate cancer susceptibility in hereditary prostate cancer (Smith *et al. Science 274*:1371-1374, 1996 and Berthon *et al. Am. J. Hum. Genet. 62*:1416-1424, 1998). These results suggest that P501S may play a role in prostate cancer malignancy.

From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for the purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the present invention is not limited except as by the appended claims.

CLAIMS

1. An isolated polypeptide comprising at least an immunogenic portion of a prostate-specific protein, or a variant thereof, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) sequences recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535 and 536;

(b) sequences that hybridize to any of the foregoing sequences under moderately stringent conditions; and

(c) complements of any of the sequence of (a) or (b).

2. An isolated polypeptide according to claim 1, wherein the polypeptide comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID No: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535 and 536, or a complement of any of the foregoing polynucleotide sequences.

3. An isolated polypeptide comprising a sequence recited in any one of SEQ ID NO: 108, 112, 113, 114, 172, 176, 178, 327, 329, 331, 339, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534 and 537-550.

4. An isolated polynucleotide encoding at least 15 contiguous amino acid residues of a prostate-specific protein, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535 and 536, or a complement of any of the foregoing sequences.

5. An isolated polynucleotide encoding a prostate-specific protein, or a variant thereof, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535 and 536, or a complement of any of the foregoing sequences.

6. An isolated polynucleotide comprising a sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535 and 536.

7. An isolated polynucleotide comprising a sequence that hybridizes under moderately stringent conditions to a sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535 and 536.

8. An isolated polynucleotide complementary to a polynucleotide according to any one of claims 4-7.

9. An expression vector comprising a polynucleotide according to any one of claims 4-8.

10. A host cell transformed or transfected with an expression vector according to claim 9.

11. An isolated antibody, or antigen-binding fragment thereof, that specifically binds to a prostate-specific protein, the protein comprising an amino acid sequence encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535 and 536 or a complement of any of the foregoing polynucleotide sequences.

12. A monoclonal antibody that specifically binds to an amino acid sequence selected from the group consisting of SEQ ID NO: 496, 504, 505, 509-517, 519, 520, 522 and 539-551.

13. A monoclonal antibody comprising a complementarity determining region selected from the group consisting of SEQ ID NO: 502, 503 and 506-508.

14. A fusion protein comprising at least one polypeptide according to claim 1.

15. A fusion protein according to claim 14, wherein the fusion protein comprises an expression enhancer that increases expression of the fusion protein in a host cell transfected with a polynucleotide encoding the fusion protein.

16. A fusion protein according to claim 14, wherein the fusion protein comprises a T helper epitope that is not present within the polypeptide of claim 1.

17. A fusion protein according to claim 14, wherein the fusion protein comprises an affinity tag.

18. An isolated polynucleotide encoding a fusion protein according to claim 14.

19.. A pharmaceutical composition comprising a physiologically acceptable carrier and at least one component selected from the group consisting of:

(a) a polypeptide according to claim 1;

(b) a polynucleotide according to claim 4;

(c) an antibody according to any one of claims 11-13;

(d) a fusion protein according to claim 14; and

(e) a polynucleotide according to claim 18.

20. A vaccine comprising an immunostimulant and at least one component selected from the group consisting of:

(a) a polypeptide according to claim 1;

(b) a polynucleotide according to claim 4;

(c) an antibody according to any one of claims 11-13;

(d) a fusion protein according to claim 14; and

(e) a polynucleotide according to claim 18.

21. A vaccine according to claim 20, wherein the immunostimulant is an adjuvant.

22. A vaccine according to claim 20, wherein the immunostimulant induces a predominantly Type I response.

23. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a pharmaceutical composition according to claim 19.

24. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a vaccine according to claim 20.

25. A pharmaceutical composition comprising an antigen-presenting cell that expresses a polypeptide according to claim 1, in combination with a pharmaceutically acceptable carrier or excipient.

26. A pharmaceutical composition according to claim 25, wherein the antigen presenting cell is a dendritic cell or a macrophage.

27. A vaccine comprising an antigen-presenting cell that expresses a polypeptide according to claim 1, in combination with an immunostimulant.

28. A vaccine according to claim 27, wherein the immunostimulant is an adjuvant.

29. A vaccine according to claim 27, wherein the immunostimulant induces a predominantly Type I response.

30. A vaccine according to claim 27, wherein the antigen-presenting cell is a dendritic cell.

31. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of an antigen-presenting cell that expresses a polypeptide encoded by a polynucleotide recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535 and 536, and thereby inhibiting the development of a cancer in the patient.

32. A method according to claim 31, wherein the antigen-presenting cell is a dendritic cell.

33. A method according to any one of claims 23, 24 and 31, wherein the cancer is prostate cancer.

34. A method for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a prostate-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535 and 536; and

(ii) complements of the foregoing polynucleotides;

wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the prostate-specific protein from the sample.

35. A method according to claim 34, wherein the biological sample is blood or a fraction thereof.

---

36. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated according to the method of claim 50.

37. A method for stimulating and/or expanding T cells specific for a prostate-specific protein, comprising contacting T cells with at least one component selected from the group consisting of:

(i) a polypeptide according to claim 1;

(ii) a polypeptide encoded by a polynucleotide comprising a sequence provided in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535 and 536;

(iii) a polynucleotide encoding a polypeptide of (i) or (ii); and

(iv) an antigen presenting cell that expresses a polypeptide of (i) or (ii),

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

38. An isolated T cell population, comprising T cells prepared according to the method of claim 37.

39. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population according to claim 38.

40. A method for inhibiting the development of a cancer in a patient, comprising the steps of:

(a) incubating $CD4^+$ and/or CD8+ T cells isolated from a patient with at least one component selected from the group consisting of:

(i) a polypeptide according to claim 1;

(ii) a polypeptide encoded by a polynucleotide comprising a sequence of any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535 and 536;

(iii) a polynucleotide encoding a polypeptide of (i) or (ii); or

(iv) an antigen-presenting cell that expresses a polypeptide of (i) or (ii);

such that T cells proliferate; and

(b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient.

41. A method for inhibiting the development of a cancer in a patient, comprising the steps of:

(a) incubating $CD4^+$ and/or CD8+ T cells isolated from a patient with at least one component selected from the group consisting of:

(i) a polypeptide according to claim 1;

(ii) a polypeptide encoded by a polynucleotide comprising a sequence of any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535 and 536;

(iii) a polynucleotide encoding a polypeptide of (i) or (ii); or

(iv) an antigen-presenting cell that expresses a polypeptide of (i) or (ii);

such that T cells proliferate;

(b) cloning at least one proliferated cell to provide cloned T cells; and

(c) administering to the patient an effective amount of the cloned T cells, and thereby inhibiting the development of a cancer in the patient.

42. A method for determining the presence or absence of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with a binding agent that binds to a prostate-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535 and 536; and

(ii) complements of the foregoing polynucleotides;

(b) detecting in the sample an amount of polypeptide that binds to the binding agent; and

(c) comparing the amount of polypeptide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

43. A method according to claim 42, wherein the binding agent is an antibody.

44. A method according to claim 43, wherein the antibody is a monoclonal antibody.

45. A method according to claim 42, wherein the cancer is prostate cancer.

46. A method for monitoring the progression of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a prostate-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence of any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535 and 536, or a complement of any of the foregoing polynucleotides;

(b) detecting in the sample an amount of polypeptide that binds to the binding agent;

(c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

(d) comparing the amount of polypeptide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

47. A method according to claim 46, wherein the binding agent is an antibody.

48. A method according to claim 47, wherein the antibody is a monoclonal antibody.

49. A method according to claim 46, wherein the cancer is a prostate cancer.

50. A method for determining the presence or absence of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a prostate-specific protein,

wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence of any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535 and 536, or a complement of any of the foregoing polynucleotides;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and

(c) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

51. A method according to claim 50, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.

52. A method according to claim 50, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

53. A method for monitoring the progression of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a prostate-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence of any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535 and 536, or a complement of any of the foregoing polynucleotides;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide;

(c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

(d) comparing the amount of polynucleotide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

54. A method according to claim 53, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.

55. A method according to claim 53, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

56. A diagnostic kit, comprising:

(a) one or more antibodies according to claim 11; and

(b) a detection reagent comprising a reporter group.

57. A kit according to claim 56, wherein the antibodies are immobilized on a solid support.

58. A kit according to claim 56, wherein the detection reagent comprises an anti-immunoglobulin, protein G, protein A or lectin.

59. A kit according to claim 56, wherein the reporter group is selected from the group consisting of radioisotopes, fluorescent groups, luminescent groups, enzymes, biotin and dye particles.

60. An oligonucleotide comprising 10 to 40 contiguous nucleotides that hybridize under moderately stringent conditions to a polynucleotide that encodes a prostate-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45,

47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-476, 524, 526, 530, 531, 533, 535 and 536, or a complement of any of the foregoing polynucleotides.

61. A oligonucleotide according to claim 60, wherein the oligonucleotide comprises 10-40 contiguous nucleotides recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-476, 524, 526, 530, 531, 533, 535 and 536.

62. A diagnostic kit, comprising:

(a) an oligonucleotide according to claim 61; and

(b) a diagnostic reagent for use in a polymerase chain reaction or hybridization assay.

63. A host cell according to claim 10, wherein the cell is selected from the group consisting of: *E. coli*, baculovirus and mammalian cells.

64. A recombinant protein produced by a host cell according to claim 10.

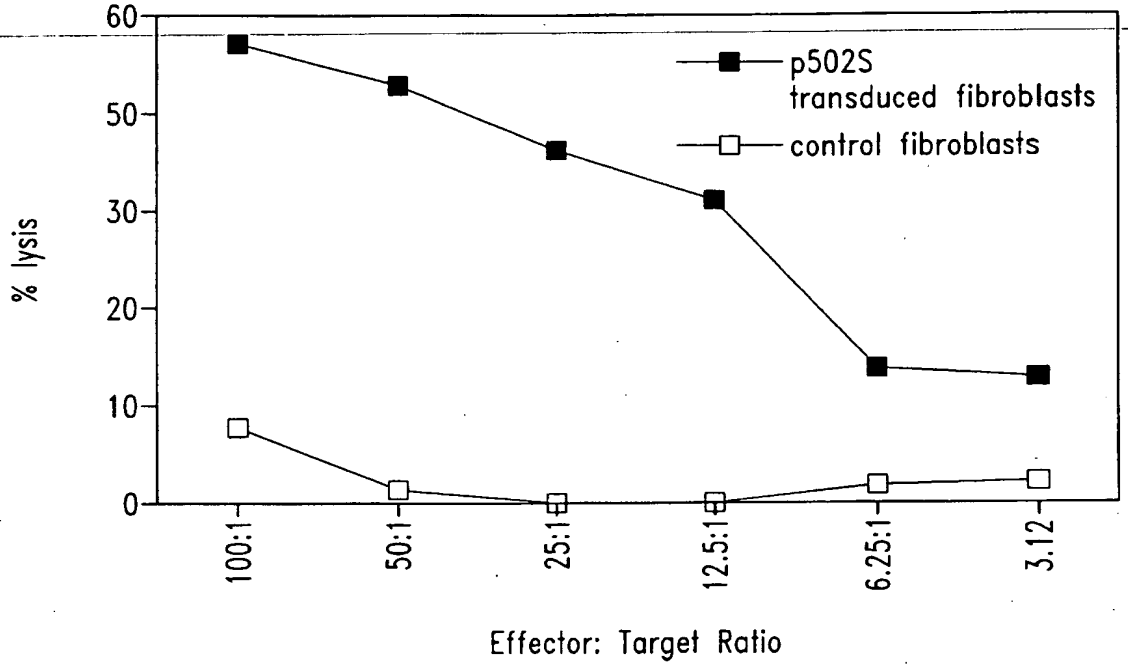


Fig. 1

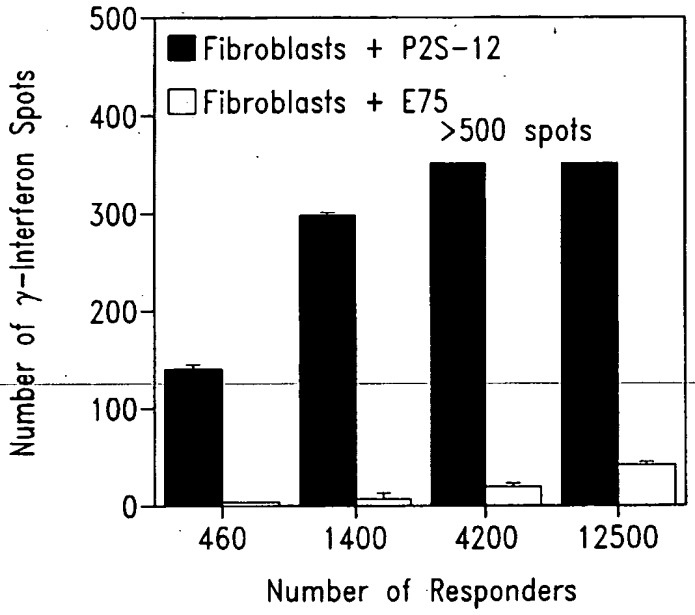


Fig. 2A

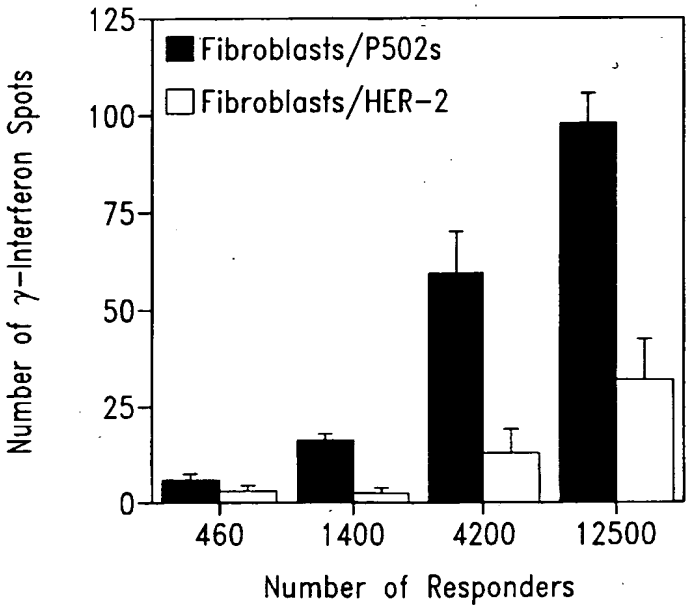


Fig. 2B

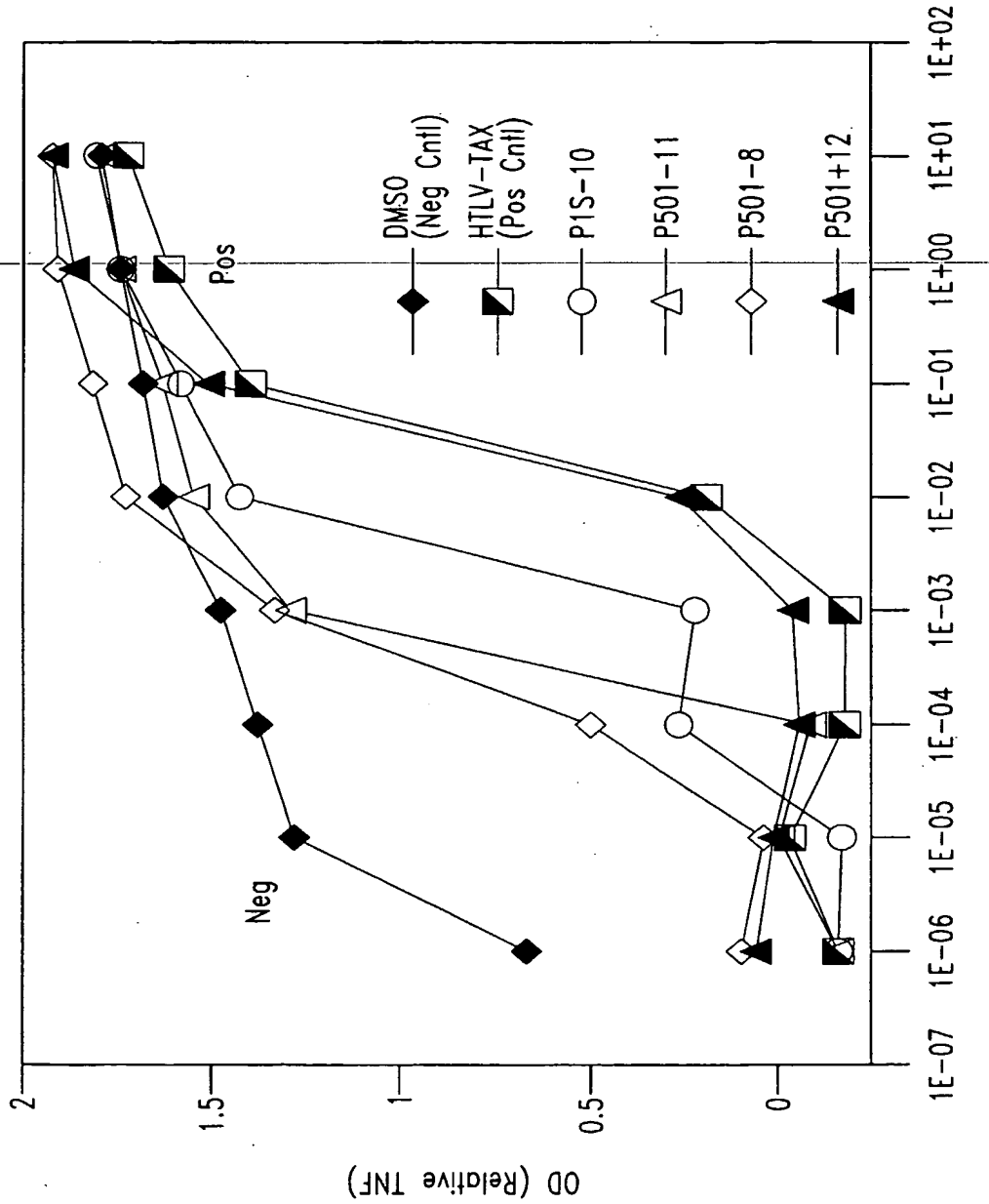


Fig. 3

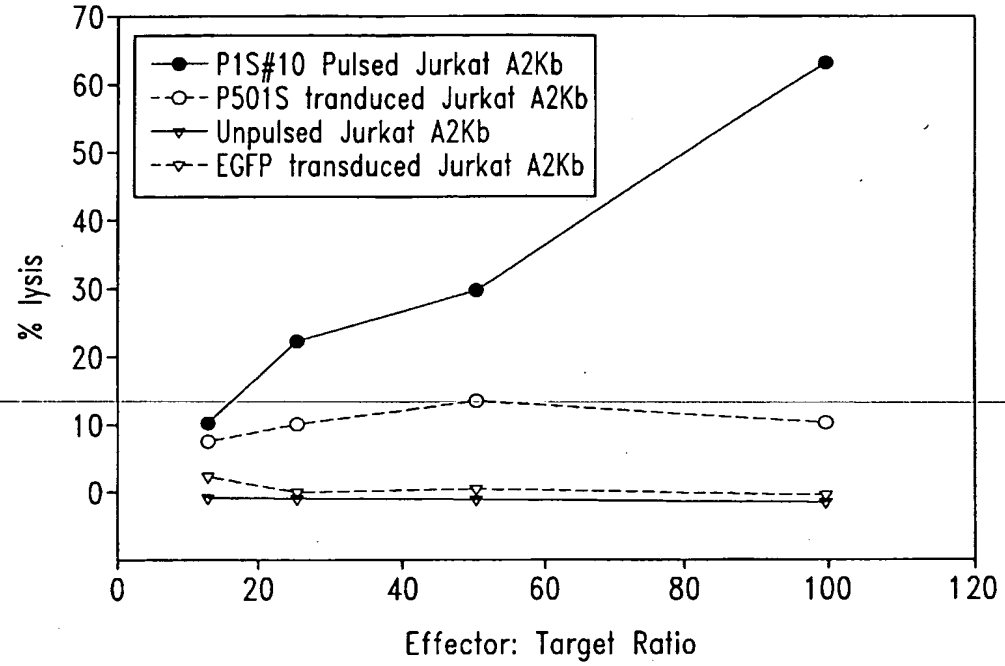


Fig. 4

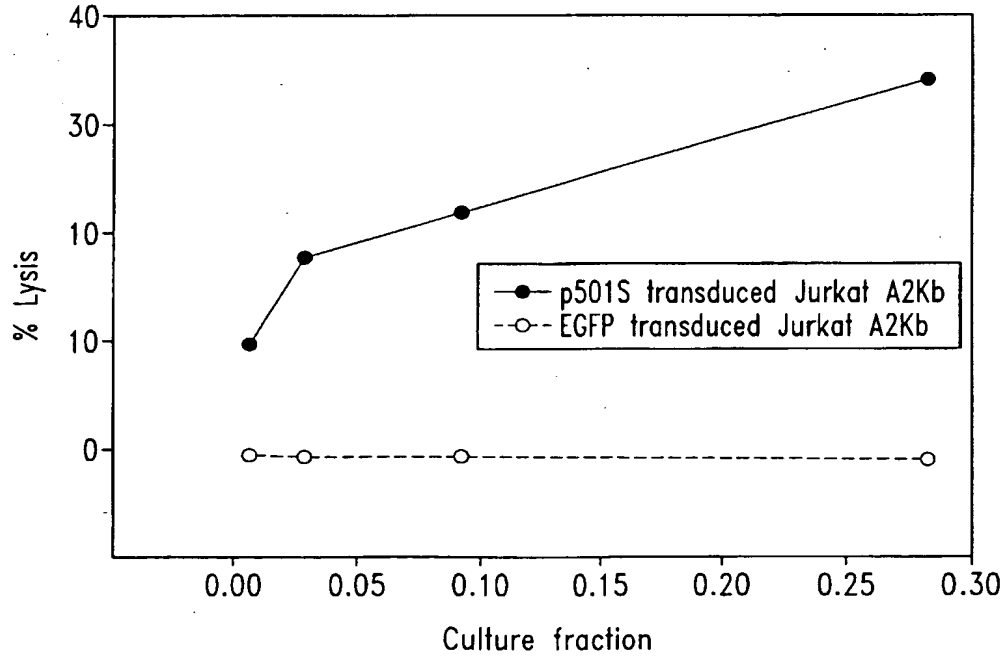


Fig. 5

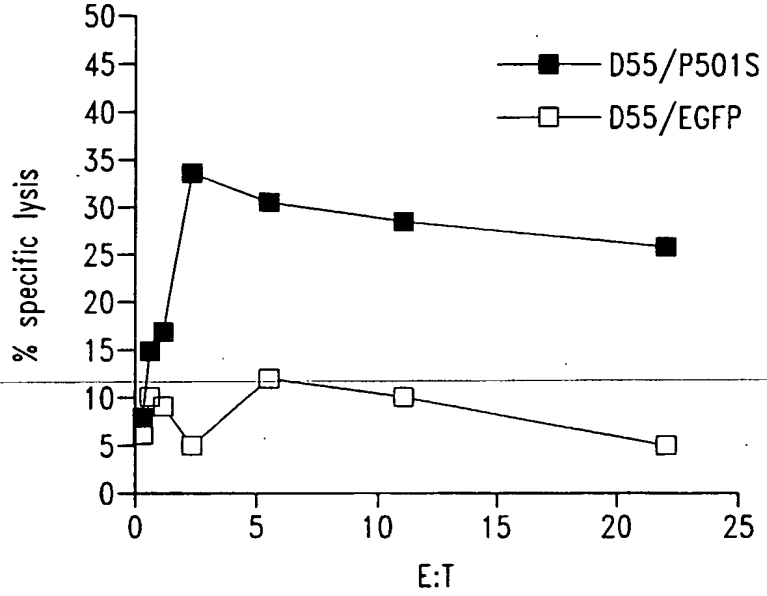


Fig. 6A

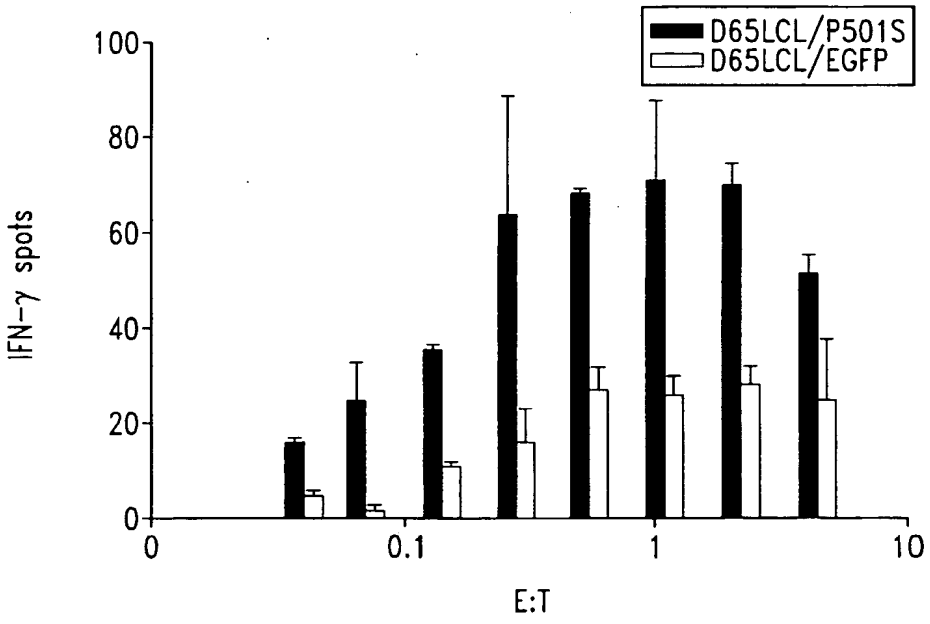
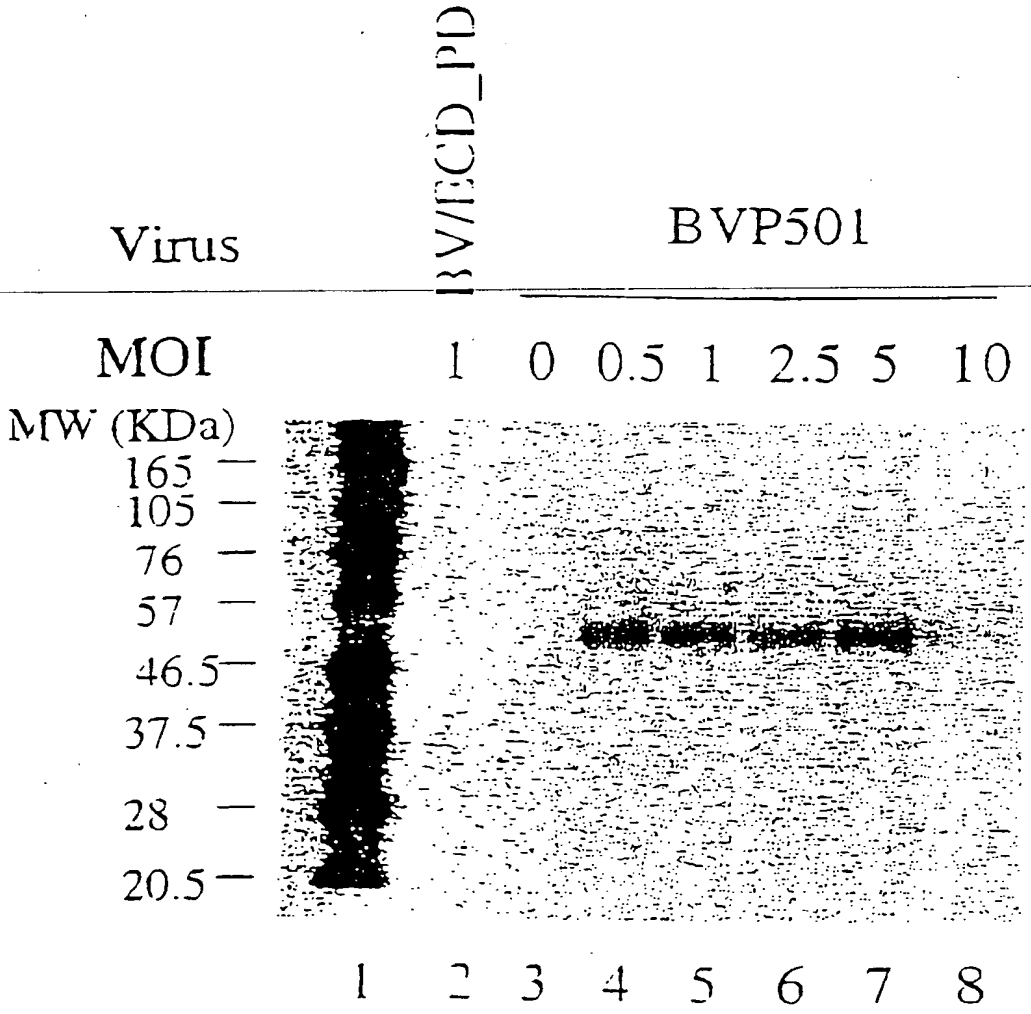


Fig. 6B

Expression of P501S by the Baculovirus Expression System



0.6 million high 5 cells in 6-well plate were infected with an unrelated control virus BV/ECD_PD (lane 2), without virus (lane 3), or with recombinant baculovirus for P501 at different MOIs (lane 4 – 8). Cell lysates were run on SDS-PAGE under the reducing conditions and analyzed by Western blot with a monoclonal antibody against P501S (P501S-10E3-G4D3). Lane 1 is the biotinylated protein molecular weight marker (Stratats).

Fig. 7

Figure 8. Mapping of the epitope recognized by 10E3-G4-D3

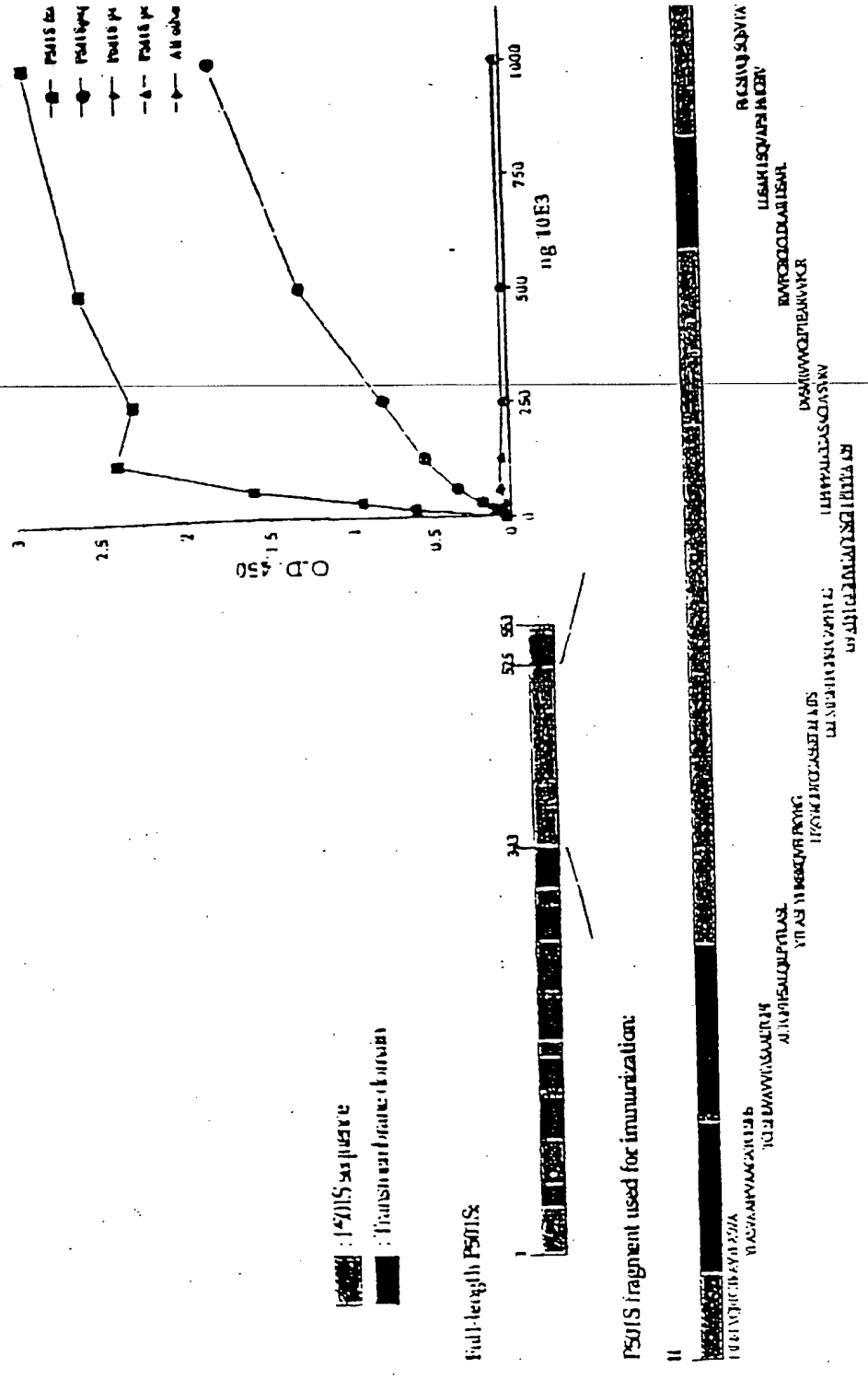


Fig. 8

Schematic of P501S with predicted
transmembrane, cytoplasmic, and extracellular regions

MVQRLWVSRLLRHRK AQLLLVNLLTFGLEVCLAAGIT **YVPPLLLEVGVEEKFM**
TMVLGIGPVLGLVCYPLLGSAS

DHWRGRYGRRRP FIWALSLGILLSLFLIPRAGWL **AGLLCPDPRPLE** LALLILGVGLLDFCGQVCFTPL

EALLSDLFRDPDHCRQ AYSVYAFMISLGGCLGYLLPAI **DWDTSALAPYLGTQEE**

CLFGLLTLIFLTCVAATLLV *AEEAALGPTEPAEGLSAPSLSPHCCPCRARLAFRNLGALLPRL*

HQLCCRMPRTLRR LFVAELCSWMALMTFTLFYTDF **VGEGLYQGVPRAEPGTEARRHYDEGVR**

MGSLGLFLQCAISLVFSLVM *DRLVQRFGTRAVYLAS* VAAFPVAAGATCLSHSVAVVTA **SAA**

LTGFTFSALQILPYTLASLY *HREKQVFLPKYRGDTGGASSEDSLMTSFLPGPKPGAPFPNGHVGAGGSGL*

LPPPPALCGASACDVSVRVVVGEPTEARVVPGRG ICLDLAILDSAFLLSQVAPSLF **MGSIVQLSQS**

VTAYMVSAAGLGLVAIYFAT *QVVFDKSDLAKYSA*

Underlined sequence: Predicted transmembrane domain; **Bold sequence**: Predicted extracellular domain; *Italic sequence*: Predicted intracellular domain. Sequence in bold/underlined: used generate polyclonal rabbit serum

Localization of domains predicted using HMMTOP (G.E. Tusnady an I. Simon (1998) Principles Governing Amino Acid Composition of Integral Membrane Proteins: Applications to topology Prediction.J.Mol Biol. 283, 489-506.

Fig. 9

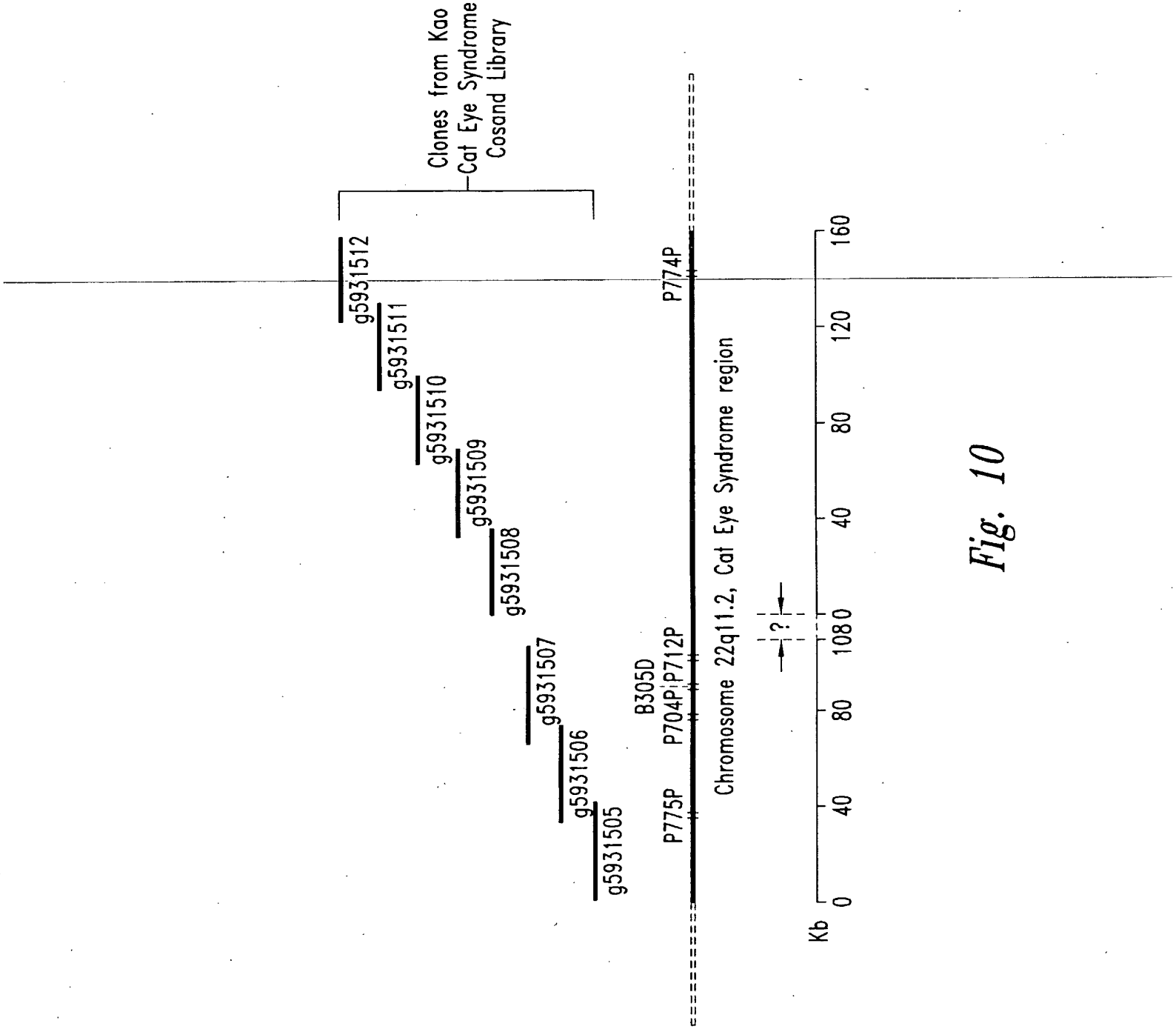


Fig. 10

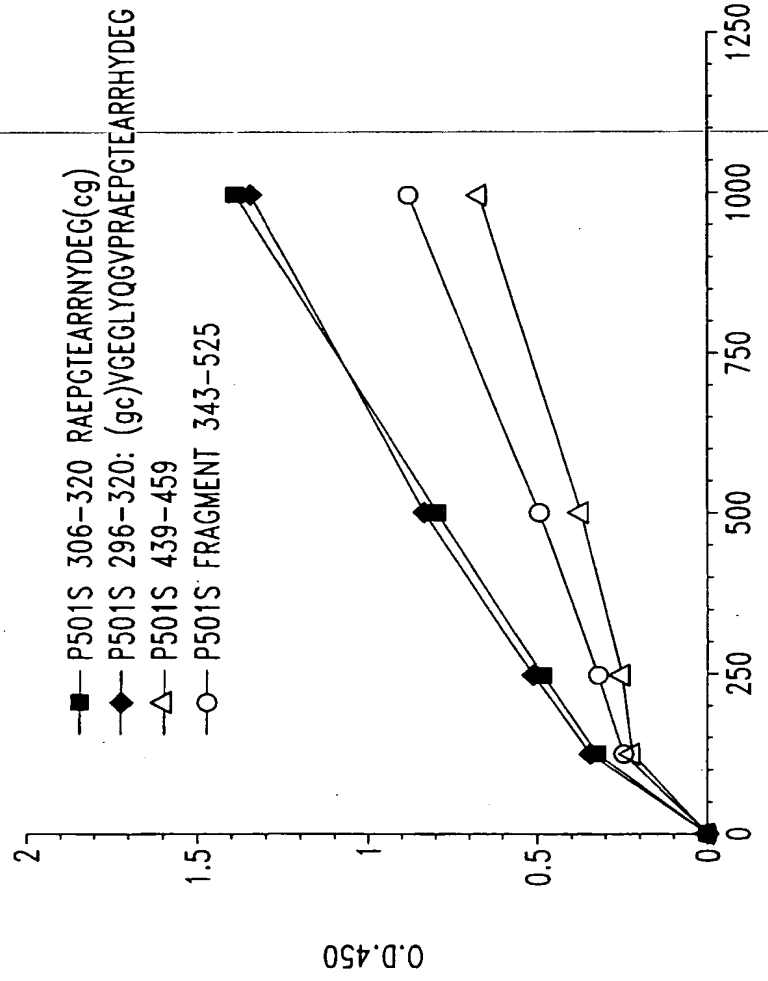


Fig. 11

SEQUENCE LISTING

<110> Corixa Corporation
Xu, Jiangchun
Dillon, Davin C.
Mitcham, Jennifer L.
Harlocker, Susan Louise
Jiang Yuqui
Reed, Steven G.
Kalos, Michael
Fanger, Gary
Retter, Mark
Solk, John
Day, Craig
Skeiky, Yasir A.W.
Wang, Aijun

<120> COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF PROSTATE CANCER

<130> 210121.42720PC

<140> PCT
<141> 2000-11-09

<160> 551

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 814
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(814)
<223> n = A,T,C or G

<400> 1
```
ttttttttttt tttttcacag tataacagct ctttatttct gtgagttcta ctaggaaatc      60
atcaaatctg agggttgtct ggaggacttc aatacacctc cccccatagt gaatcagctt     120
ccaggggtc cagtccctct ccttacttca tccccatccc atgccaaagg aagaccctcc      180
ctccttggct cacagccttc tctaggcttc ccagtgcctc caggacagag tgggttatgt     240
tttcagctcc atccttgctg tgagtgtctg gtgcgttgtg cctccagctt ctgctcagtg     300
cttcatggac agtgtccagc acatgtcact ctccactctc tcagtgtgga tccactagtt     360
ctagagcggc cgccaccgcg gtggagctcc agcttttgtt ccctttagtg agggttaatt     420
gcgcgcttgg cgtaatcatg gtcataactg tttcctgtgt gaaattgtta tccgctcaca     480
attccacaca acatacgagc cggaagcata aagtgtaaag cctggggtgc ctaatgagtg     540
anctaactca cattaattgc gttgcgctca ctgnccgctt tccagtcngg aaaactgtcg     600
tgccagctgc attaatgaat cggccaacgc ncggggaaaa gcggtttgcg ttttgggggc     660
tcttccgctt ctcgctcact nantcctgcg ctcggtcntt cggctgcggg gaacggtatc     720
actcctcaaa ggnggtatta cggttatccn naaatcnggg gatacccngg aaaaaanttt     780
aacaaaaggg cancaaaggg cngaaacgta aaaa                                 814
```

<210> 2
<211> 816
<212> DNA

<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(816)
<223> n = A,T,C or G

<400> 2
acagaaatgt tggatggtgg agcacctttc tatacgactt acaggacagc agatggggaa 60
ttcatggctg ttggagcaat agaaccccag ttctacgagc tgctgatcaa aggacttgga 120
ctaaagtctg atgaacttcc caatcagatg agcatggatg attggccaga aatgaagaag 180
aagtttgcag atgtatttgc aaagaagacg aaggcagagt ggtgtcaaat ctttgacggc 240
acagatgcct gtgtgactcc ggttctgact tttgaggagg ttgttcatca tgatcacaac 300
aaggaacggg gctcgtttat caccagtgag gagcaggacg tgagcccccg ccctgcacct 360
ctgctgttaa acaccccagc catcccttct ttcaaaaggg atccactagt tctagaagcg 420
gccgccaccg cggtggagct ccagcttttg ttccctttag tgagggttaa ttgcgcgctt 480
ggcgtaatca tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccccc 540
aacatacgag ccggaacata aagtgttaag cctggggtgc ctaatgantg agctaactcn 600
cattaattgc gttgcgctca ctgcccgctt tccagtcggg aaaactgtcg tgccactgcn 660
ttantgaatc ngccaccccc cgggaaaagg cggttgcntt ttgggcctct tccgctttcc 720
tcgctcattg atcctngcnc ccggtcttcg gctgcggnga acggttcact cctcaaaggc 780
ggtntnccgg ttatccccaa acnggggata cccnga 816

<210> 3
<211> 773
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(773)
<223> n = A,T,C or G

<400> 3
cttttgaaag aagggatggc tggggtgttt aacagcagag gtgcagggcg ggggctcacg 60
tcctgctcct cactggtgat aaacgagccc cgttccttgt tgtgatcatg atgaacaacc 120
tcctcaaaag tcagaaccgg agtcacacag gcatctgtgc cgtcaaagat ttgacaccac 180
tctgccttcg tcttctttgc aaatacatct gcaaacttct tcttcatttc tggccaatca 240
tccatgctca tctgattggg aagttcatca gactttagtc canntccttt gatcagcagc 300
tcgtagaact ggggttctat tgctccaaca gccatgaatt ccccatctgc tgtcctgtaa 360
gtcgtataga aaggtgctcc accatccaac atgttctgtc ctcgaggggg ggcccggtac 420
ccaattcgcc ctatantgag tcgtattacg cgcgctcact ggccgtcgtt ttacaacgtc 480
gtgactggga aaaccctggg cgttaccaac ttaatcgcct tgcagcacat ccccctttcg 540
ccagctgggc gtaatancga aaaggcccgc accgatcgcc cttccaacag ttgcgcacct 600
gaatgggnaa atgggacccc cctgttaccg cgcattnaac ccccgcnggg tttngttgtt 660
acccccacnt nnaccgctta cactttgcca gcgccttanc gcccgctccc tttcnccttt 720
cttcccttcc tttcncnccn ctttcccccg gggtttcccc cntcaaaccc cna 773

<210> 4
<211> 828
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(828)
<223> n = A,T,C or G

<400> 4
cctcctgagt cctactgacc tgtgctttct ggtgtggagt ccagggctgc taggaaaagg 60
aatgggcaga cacaggtgta tgccaatgtt tctgaaatgg gtataatttc gtcctctcct 120
tcggaacact ggctgtctct gaagacttct cgctcagttt cagtgaggac acacacaaag 180
acgtgggtga ccatgttgtt tgtggggtgc agagatggga ggggtggggc ccaccctgga 240
agagtggaca gtgacacaag gtggacactc tctacagatc actgaggata agctggagcc 300
acaatgcatg aggcacacac acagcaagga tgacnctgta aacatagccc acgctgtcct 360
gnggggcactg ggaagcctan atnaggccgt gagcanaaag aaggggagga tccactagtt 420
ctanagcggc cgccaccgcg gtgganctcc ancttttgtt ccctttagtg agggttaatt 480
gcgcgcttgg cntaatcatg gtcatanctn tttcctgtgt gaaattgtta tccgctcaca 540
attccacaca acatacganc cggaaacata aantgtaaac ctggggtgcc taatgantga 600
ctaactcaca ttaattgcgt tgcgctcact gcccgctttc caatcnggaa acctgtcttg 660
ccncttgcat tnatgaatcn gccaaccccc ggggaaaagc gtttgcgttt tgggcgctct 720
tccgcttcct cnctcantta ntccctncnc tcggtcattc cggctgcngc aaaccggttc 780
accncctcca aaggggtat tccggtttcc ccnaatccgg gganancc 828

<210> 5
<211> 834
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(834)
<223> n = A,T,C or G

<400> 5
ttttttttt tttttactga tagatggaat ttattaagct tttcacatgt gatagcacat 60
agttttaatt gcatccaaag tactaacaaa aactctagca atcaagaatg gcagcatgtt 120
attttataac aatcaacacc tgtggctttt aaaatttggt tttcataaga taatttatac 180
tgaagtaaat ctagccatgc ttttaaaaaa tgctttaggt cactccaagc ttggcagtta 240
acatttggca taaacaataa taaaacaatc acaatttaat aaataacaaa tacaacattg 300
taggccataa tcatatacag tataaggaaa aggtggtagt gttgagtaag cagttattag 360
aatagaatac cttggcctct atgcaaatat gtctagacac tttgattcac tcagccctga 420
cattcagttt tcaaagtagg agacaggttc tacagtatca ttttacagtt tccaacacat 480
tgaaaacaag tagaaaatga tgagttgatt tttattaatg cattacatcc tcaagagtta 540
tcaccaaccc ctcagttata aaaaattttc aagttatatt agtcatataa cttggtgtgc 600
ttattttaaa ttagtgctaa atggattaag tgaagacaac aatggtcccc taatgtgatt 660
gatattggtc atttttacca gcttctaaat ctnaactttc aggcttttga actggaacat 720
tgnatnacag tgttccanag ttncaaccta ctggaacatt acagtgtgct tgattcaaaa 780
tgttattttg ttaaaaatta aattttaacc tggtggaaaa ataatttgaa atna 834

<210> 6
<211> 818
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(818)
<223> n = A,T,C or G

<400> 6
ttttttttt ttttttttt aagaccctca tcaatagatg gagacataca gaaatagtca 60
aaccacatct acaaaatgcc agtatcaggc ggcggcttcg aagccaaagt gatgtttgga 120
tgtaaagtga aatattagtt ggcggatgaa gcagatagtg aggaaagttg agccaataat 180
gacgtgaagt ccgtggaagc ctgtggctac aaaaaatgtt gagccgtaga tgccgtcgga 240
aatggtgaag ggagactcga agtactctga ggcttgtagg agggtaaaat agagacccag 300

taaaattgta ataagcagtg cttgaattat ttggtttcgg ttgttttcta ttagactatg 360
gtgagctcag gtgattgata ctcctgatgc gagtaatacg gatgtgttta ggagtgggac 420
ttctagggga tttagcgggg tgatgcctgt tgggggccag tgccctccta gttggggggt 480
aggggctagg ctggagtggt aaaaggctca gaaaaatcct gcgaagaaaa aaacttctga 540
ggtaataaat aggattatcc cgtatcgaag gccttttggg acaggtggtg tgtggtggcc 600
ttggtatgtg ctttctcgtg ttacatcgcg ccatcattgg tatatggtta gtgtgttggg 660
ttantanggc ctantatgaa gaacttttgg antggaatta aatcaatngc ttggccggaa 720
gtcattanga nggctnaaaa ggccctgtta ngggtctggg ctnggtttta cccnacccat 780
ggaatncncc ccccggacna ntgnatccct attcttaa 818

<210> 7
<211> 817
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(817)
<223> n = A,T,C or G

<400> 7
ttttttttttt ttttttttttt tggctctaga gggggtagag ggggtgctat agggtaaata 60
cgggccctat ttcaaagatt tttaggggaa ttaattctag gacgatgggt atgaaactgt 120
ggtttgctcc acagatttca gagcattgac cgtagtatac ccccggtcgt gtagcggtga 180
aagtggtttg gtttagacgt ccgggaattg catctgtttt taagcctaat gtggggacag 240
ctcatgagtg caagacgtct tgtgatgtaa ttattatacn aatgggggct tcaatcggga 300
gtactactcg attgtcaacg tcaaggagtc gcaggtcgcc tggttctagg aataatgggg 360
gaagtatgta ggaattgaag attaatccgc cgtagtcggt gttctcctag gttcaatacc 420
attggtggcc aattgatttg atggtaaggg gagggatcgt tgaactcgtc tgttatgtaa 480
aggatncctt ngggatggga aggcnatnaa ggactangga tnaatggcgg gcangatatt 540
tcaaacngtc tctanttcct gaaacgtctg aaatgttaat aanaattaan tttngttatt 600
gaatnttnng gaaaagggct tacaggacta gaaaccaaat angaaaanta atnntaangg 660
cnttatcntn aaaggtnata accnctccta tnatcccacc caatngnatt ccccacncnn 720
acnattggat nccccanttc canaaanggc cnccccccgg tgnanncenc cttttgttcc 780
cttnantgan ggttattcnc ccctngcntt atcancc 817

<210> 8
<211> 799
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(799)
<223> n = A,T,C or G

<400> 8
catttccggg tttactttct aaggaaagcc gagcggaagc tgctaacgtg ggaatcggtg 60
cataaggaga actttctgct ggcacgcgct agggacaagc gggagagcga ctccgagcgt 120
ctgaagcgca cgtcccagaa ggtggacttg gcactgaaac agctgggaca catccgcgag 180
tacgaacagc gcctgaaagt gctggagcgg gaggtccagc agtgtagccg cgtcctgggg 240
tgggtggccg angcctganc cgctctgcct tgctgccccc angtgggccg ccacccccctg 300
acctgcctgg gtccaaacac tgagccctgc tggcggactt caagganaac ccccacangg 360
ggattttgct cctanantaa ggctcatctg ggcctcggcc cccccacctg gttggccttg 420
tctttgangt gagccccatg tccatctggg ccactgtcng gaccaccttt ngggagtgtt 480
ctccttacaa ccacannatg cccggctcct cccggaaacc antcccancc tgngaaggat 540
caagncctgn atccactnnt nctanaaccg gccnccnccg cngtggaacc cncctttntgt 600
tccttttcnt tnagggttaa tnncgccttg gccttnccan ngtcctncnc nttttccnnt 660

gttnaaattg ttangcnccc nccnntcccn cnncnncnan cccgacccnn annttnnann 720
ncctgggggt nccnncngat tgacccnncc nccctntant tgcnttnggg nncnntgccc 780
ctttccctct ngggannng 799

<210> 9
<211> 801
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(801)
<223> n = A,T,C or G

<400> 9
acgccttgat cctcccaggc tgggactggt tctgggagga gccgggcatg ctgtggtttg 60
taangatgac actcccaaag gtggtcctga cagtggccca gatggacatg gggctcacct 120
caaggacaag gccaccaggt gcgggggccg aagcccacat gatccttact ctatgagcaa 180
aatcccctgt gggggcttct ccttgaagtc cgccancagg gctcagtctt tggacccang 240
caggtcatgg ggttgtngnc caactggggg ccncaacgca aaanggcnca gggcctcngn 300
cacccatccc angacgcggc tacactnctg gacctcccnc tccaccactt tcatgcgctg 360
ttcntacccg cgnatntgtc ccanctgttt cngtgccnac tccanctttct nggacgtgcg 420
ctacatacgc ccggantcnc nctcccgctt tgtccctatc cacgtnccan caacaaattt 480
cnccntantg caccnattcc cacntttnnc agntttccnc nncgngcttc cttntaaaag 540
ggttganccc cggaaaatnc cccaaagggg gggggccngg tacccaactn ccccctnata 600
gctgaantcc ccatnaccnn gnctcnatgg anccntccnt tttaannacn ttctnaactt 660
gggaanancc ctcgnccntn cccccnttaa tcccnccttg cnangnncnt cccccnntcc 720
ncccnnntng gcntntnann cnaaaaaggc ccnnnancaa tctcctnncn cctcanttcg 780
ccancctcg aaatcggccn c 801

<210> 10
<211> 789
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(789)
<223> n = A,T,C or G

<400> 10
cagtctatnt ggccagtgtg gcagctttcc ctgtggctgc cggtgccaca tgcctgtccc 60
acagtgtggc cgtggtgaca gcttcagccg ccctcaccgg gttcaccttc tcagccctgc 120
agatcctgcc ctacacactg gcctccctct accaccggga gaagcaggtg ttcctgccca 180
aataccgagg ggacactgga ggtgctagca gtgaggacag cctgatgacc agcttcctgc 240
caggccctaa gcctggagct cccttcccta atggacacgt gggtgctgga ggcagtggcc 300
tgctcccacc tccacccgcg ctctgcgggg cctctgcctg tgatgtctcc gtacgtgtgg 360
tggtgggtga gcccaccgan gccagggtgg ttccgggccg gggcatctgc ctggacctcg 420
ccatcctgga tagtgcttcc tgctgtccca ngtggcccca tccctgttta tgggctccat 480
tgtccagctc agccagtctg tcactgccta tatggtgtct gccgcaggcc tgggtctggt 540
cccatttact ttgctacaca ggtantattt gacaagaacg anttggccaa atactcagcg 600
ttaaaaaatt ccagcaacat tgggggtgga aggcctgcct cactgggtcc aactccccgc 660
tcctgttaac cccatggggc tgccggcttg gccgccaatt tctgttgctg ccaaantnat 720
gtggctctct gctgccacct gttgctggct gaagtgcnta cngcncanct ngggggggtng 780
ggngttccc 789

<210> 11
<211> 772

<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(772)
<223> n = A,T,C or G

<400> 11
cccaccctac ccaaatatta gacaccaaca cagaaaagct agcaatggat tcccttctac 60
tttgttaaat aaataagtta aatatttaaa tgcctgtgtc tctgtgatgg caacagaagg 120
accaacaggc cacatcctga taaaaggtaa gaggggggtg gatcagcaaa aagacagtgc 180
tgtgggctga ggggacctgg ttcttgtgtg ttgcccctca ggactcttcc cctacaaata 240
actttcatat gttcaaatcc catggaggag tgtttcatcc tagaaactcc catgcaagag 300
ctacattaaa cgaagctgca ggttaagggg cttanagatg ggaaaccagg tgactgagtt 360
tattcagctc ccaaaaaccc ttctctaggt gtgtctcaac taggaggcta gctgttaacc 420
ctgagcctgg gtaatccacc tgcagagtcc ccgcattcca gtgcatggaa cccttctggc 480
ctccctgtat aagtccagac tgaaaccccc ttggaaggnc tccagtcagg cagccctana 540
aactggggaa aaaagaaaag gacgccccan cccccagctg tgcanctacg cacctcaaca 600
gcacagggtg gcagcaaaaa aaccacttta ctttggcaca aacaaaaact ngggggggca 660
accccggcac cccnanggggg gttaacagga ancngggnaa cntggaaccc aattnaggca 720
ggcccnccac cccnaatntt gctgggaaat ttttcctccc ctaaattntt tc 772

<210> 12
<211> 751
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(751)
<223> n = A,T,C or G

<400> 12
gccccaattc cagctgccac accacccacg gtgactgcat tagttcggat gtcatacaaa 60
agctgattga agcaaccctc tacttttttgg tcgtgagcct tttgcttggt gcaggtttca 120
ttggctgtgt tggtgacgtt gtcattgcaa cagaatgggg gaaaggcact gttctctttg 180
aagtanggtg agtcctcaaa atccgtatag ttggtgaagc cacagcactt gagccctttc 240
atggtggtgt tccacacttg agtgaagtct tcctgggaac cataatcttt cttgatggca 300
ggcactacca gcaacgtcag ggaagtgctc agccattgtg gtgtacacca aggcgaccac 360
agcagctgcn acctcagcaa tgaagatgan gaggangatg aagaagaacg tcncgagggc 420
acacttgctc tcagtcttan caccatanca gcccntgaaa accaananca aagaccacna 480
cnccggctgc gatgaagaaa tnacccccncg ttgacaaact tgcatggcac tgggganccac 540
agtggcccna aaaatcttca aaaaggatgc cccatcnatt gacccccaa atgcccactg 600
ccaacagggg ctgccccacn cncnnaacga tganccnatt gnacaagatc tncntggtct 660
tnatnaacnt gaaccctgcn tngtggctcc tgttcaggnc cnnggcctga cttctnaann 720
aangaactcn gaagncccca cngganannc g 751

<210> 13
<211> 729
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(729)
<223> n = A,T,C or G

<400> 13
gagccaggcg tccctctgcc tgcccactca gtggcaacac ccgggagctg ttttgtcctt 60
tgtggancct cagcagtncc ctctttcaga actcantgcc aagancccctg aacaggagcc 120
accatgcagt gcttcagctt cattaagacc atgatgatcc tcttcaattt gctcatcttt 180
ctgtgtggtg cagccctgtt ggcagtgggc atctgggtgt caatcgatgg ggcatccttt 240
ctgaagatct tcgggccact gtcgtccagt gccatgcagt ttgtcaacgt gggctacttc 300
ctcatcgcag ccggcgttgt ggtcttagct ctaggtttcc tgggctgcta tggtgctaag 360
actgagagca agtgtgccct cgtgacgttc ttcttcatcc tcctcctcat cttcattgct 420
gaggttgcaa tgctgtggtc gccttggtgt acaccacaat ggctgagcac ttcctgacgt 480
tgctggtaat gcctgccatc aanaaaagat tatgggttcc caggaanact tcactcaagt 540
gttggaacac caccatgaaa gggctcaagt gctgtggctt cnnccaacta tacggatttt 600
gaagantcac ctacttcaaa gaaaanagtg cctttccccc atttctgttg caattgacaa 660
acgtccccaa cacagccaat tgaaaacctg cacccaaccc aaangggtcc ccaaccanaa 720
attnaaggg 729

<210> 14
<211> 816
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(816)
<223> n = A,T,C or G

<400> 14
tgctcttcct caaagttgtt cttgttgcca taacaaccac cataggtaaa gcgggcgcag 60
tgttcgctga aggggttgta gtaccagcgc gggatgctct ccttgcagag tcctgtgtct 120
ggcaggtcca cgcagtgccc tttgtcactg gggaaatgga tgcgctggag ctcgtcaaag 180
ccactcgtgt atttttcaca ggcagcctcg tccgacgcgt cggggcagtt gggggtgtct 240
tcacactcca ggaaactgtc natgcagcag ccattgctgc agcggaactg ggtgggctga 300
cangtgccag agcacactgg atggcgcctt tccatgnnan gggccctgng ggaaagtccc 360
tgancccan anctgcctct caaangcccc accttgcaca ccccgacagg ctagaatgga 420
atcttcttcc cgaaaggtag ttnttcttgt tgcccaancc ancccentaa acaaactctt 480
gcanatctgc tccgnggggg tcntantacc ancgtgggaa aagaacccca ggcngcgaac 540
caancttgtt tggatncgaa gcnataatct nctnttctgc ttggtggaca gcaccantna 600
ctgtnnanct ttagnccntg gtcctcntgg gttgnncttg aacctaatcn ccnntcaact 660
gggacaaggt aantngccnt cctttnaatt cccnancntn ccccctggtt tggggttttn 720
cncnctccta ccccagaaan nccgtgttcc cccccaacta ggggccnaaa ccnnttnttc 780
cacaaccctn ccccacccac gggttcngnt ggttng 816

<210> 15
<211> 783
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(783)
<223> n = A,T,C or G

<400> 15
ccaaggcctg ggcaggcata nacttgaagg tacaacccca ggaacccctg gtgctgaagg 60
atgtggaaaa cacagattgg cgcctactgc ggggtgacac ggatgtcagg gtagagagga 120
aagacccaaa ccaggtggaa ctgtggggac tcaaggaang cacctacctg ttccagctga 180
cagtgactag ctcagaccac ccagaggaca cggccaacgt cacagtcact gtgctgtcca 240
ccaagcagac agaagactac tgcctcgcat ccaacaangt gggtcgctgc cggggctctt 300
tcccacgctg gtactatgac cccacggagc agatctgcaa gagtttcgtt tatggaggct 360

gcttgggcaa caagaacaac taccttcggg aagaagagtg cattctancc tgtcngggtg 420
tgcaaggtgg gcctttgana ngcanctctg gggctcangc gactttcccc cagggccct 480
ccatggaaag gcgccatcca ntgttctctg gcacctgtca gcccacccag ttccgctgca 540
ncaatggctg ctgcatcnac antttcctng aattgtgaca acacccccca ntgccccaa 600
ccctcccaac aaagcttccc tgttnaaaaa tacnccantt ggcttttnac aaacncccgg 660
cncctccntt ttccccnntn aacaaagggc nctngcnttt gaactgcccn aacccnggaa 720
tctnccnngg aaaaantncc ccccctggtt cctnnaancc cctccncnaa anctnccccc 780
ccc 783

<210> 16
<211> 801
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(801)
<223> n = A,T,C or G

<400> 16
gccccaattc cagctgccac accacccacg gtgactgcat tagttcggat gtcatacaaa 60
agctgattga agcaaccctc tacttttttgg tcgtgagcct tttgcttggt gcaggtttca 120
ttggctgtgt tggtgacgtt gtcattgcaa cagaatgggg gaaaggcact gttctctttg 180
aagtagggtg agtcctcaaa atccgtatag ttggtgaagc cacagcactt gagccctttc 240
atggtggtgt tccacacttg agtgaagtct tcctgggaac cataatcttt cttgatggca 300
ggcactacca gcaacgtcag gaagtgctca gccattgtgg tgtacaccaa ggcgaccaca 360
gcagctgcaa cctcagcaat gaagatgagg aggaggatga agaagaacgt cncgagggca 420
cacttgctct ccgtcttagc accatagcag cccangaaac caagagcaaa gaccacaacg 480
ccngctgcga atgaaagaaa ntacccacgt tgacaaactg catggccact ggacgacagt 540
tggcccgaan atcttcagaa aagggatgcc ccatcgattg aacacccana tgcccactgc 600
cnacagggct gcnccncncn gaaagaatga gccattgaag aaggatcntc ntggtcttaa 660
tgaactgaaa ccntgcatgg tggcccctgt tcagggctct tggcagtgaa ttctganaaa 720
aaggaacngc ntnagccccc ccaaangana aaacaccccc gggtgttgcc ctgaattggc 780
ggccaaggan ccctgccccn g 801

<210> 17
<211> 740
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(740)
<223> n = A,T,C or G

<400> 17
gtgagagcca ggcgtccctc tgcctgccca ctcagtggca acacccggga gctgttttgt 60
cctttgtgga gcctcagcag ttccctcttt cagaactcac tgccaagagc cctgaacagg 120
agccaccatg cagtgcttca gcttcattaa gaccatgatg atcctcttca atttgctcat 180
ctttctgtgt ggtgcagccc tgttggcagt gggcatctgg gtgtcaatcg atggggcatc 240
ctttctgaag atcttcgggc cactgtcgtc cagtgccatg cagtttgtca acgtgggcta 300
cttcctcatc gcagccggcg ttgtggtctt tgctcttggt ttcctgggct gctatggtgc 360
taagacggag agcaagtgtg ccctcgtgac gttcttcttc atcctcctcc tcatcttcat 420
tgctgaagtt gcagctgctg tggtcgcctt ggtgtacacc acaatggctg aaccattcct 480
gacgttgctg gtantgcctg ccatcaanaa agattatggg ttcccaggaa aaattcactc 540
aantntggaa caccnccatg aaaagggctc caatttctgn tggcttcccc aactataccg 600
gaattttgaa agantcnccc tacttccaaa aaaaaanant tgcctttncc cccnttctgt 660
tgcaatgaaa acntcccaan acngccaatn aaaacctgcc cnnncaaaaa ggntcncaaa 720

caaaaaaant nnaagggttn 740

<210> 18
<211> 802
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(802)
<223> n = A,T,C or G

<400> 18
ccgctggttg cgctggtcca gngnagccac gaagcacgtc agcatacaca gcctcaatca 60
caaggtcttc cagctgccgc acattacgca gggcaagagc ctccagcaac actgcatatg 120
ggatacactt tactttagca gccagggtga caactgagag gtgtcgaagc ttattcttct 180
gagcctctgt tagtggagga agattccggg cttcagctaa gtagtcagcg tatgtcccat 240
aagcaaacac tgtgagcagc cggaaggtag aggcaaagtc actctcagcc agctctctaa 300
cattgggcat gtccagcagt tctccaaaca cgtagacacc agnggcctcc agcacctgat 360
ggatgagtgt ggccagcgct gccccccttgg ccgacttggc taggagcaga aattgctcct 420
ggttctgccc tgtcaccttc acttccgcac tcatcactgc actgagtgtg ggggacttgg 480
gctcaggatg tccagagacg tggttccgcc ccctcnctta atgacaccgn ccanncaacc 540
gtcggctccc gccgantgng ttcgtcgtnc ctgggtcagg gtctgctggc cnctacttgc 600
aancttcgtc nggcccatgg aattcaccnc accggaactn gtangatcca ctnnttctat 660
aaccggncgc caccgcnnnt ggaactccac tcttnttncc tttacttgag ggttaaggtc 720
acccttnncg ttaccttggt ccaaaccntn ccntgtgtcg anatngtnaa tcnggnccna 780
tnccanccnc atangaagcc ng 802

<210> 19
<211> 731
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(731)
<223> n = A,T,C or G

<400> 19
cnaagcttcc aggtnacggg ccgcnaancc tgacccnagg tancanaang cagncngcgg 60
gagcccaccg tcacgnggng gngtctttat nggagggggc ggagccacat cnctggacnt 120
cntgacccca actccccncc ncncantgca gtgatgagtg cagaactgaa ggtnacgtgg 180
caggaaccaa gancaaannc tgctccnntc caagtcggcn naggggggcgg ggctggccac 240
gcncatccnt cnagtgctgn aaagccccnn cctgtctact tgtttggaga acngcnnnga 300
catgcccagn gttanataac nggcngagag tnantttgcc tctcccttcc ggctgcgcan 360
cgngtntgct tagnggacat aacctgacta cttaactgaa cccnngaatc tnccnccccc 420
ccactaagct cagaacaaaa aacttcgaca ccactcantt gtcacctgnc tgctcaagta 480
aagtgtaccc catncccaat gtntgctnga ngctctgncc tgcnttangt tcggtcctgg 540
gaagacctat caattnaagc tatgtttctg actgcctctt gctccctgna acaancnacc 600
cnncnntcca agggggggnc ggcccccaat ccccccaacc ntnaattnan tttanccccn 660
cccccnggcc cggcctttta cnancntcnn nnacngggna aaaccnnngc tttncccaac 720
nnaatccncc t 731

<210> 20
<211> 754
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(754)
<223> n = A,T,C or G

<400> 20
tttttttttt tttttttttt taaaaacccc ctccattnaa tgnaaacttc cgaaattgtc 60
caacccccтc ntccaaatnn ccntttccgg gngggggttc caaacccaan ttannttgg 120
annttaaatt aaatnttnnt tggnggnnna anccnaatgt nangaaagtt naacccanta 180
tnancttnaa tncctggaaa ccngtngntt ccaaaaatnt ttaacсctta antccctccg 240
aaatngttna nggaaaaccc aanttctcnt aaggttgttt gaaggntnaa tnaaaanccc 300
nnccaattgt ttttngccac gcctgaatta attggnttcc gntgttttcc nttaaaanaa 360
ggnnanсccc ggttantnaa tccccccnnc cccaattata ccgantttтt ttngaattgg 420
gancccncgg gaattaacgg ggnnnntccc tnttgggggg cnggnnсccc ccccntcggg 480
ggttngggnc aggncnnaat tgtttaaggg tccgaaaaat ccctccnaga aaaaaanctc 540
ccaggntgag nntngggttt ncccccccсc canggcccct ctcgnanagt tggggtttgg 600
ggggcctggg attttntttc ccctnttncc tccсccccсc ccnggganag aggttngngt 660
tttgntcnnc ggcccсncсn aaganctttn ccganttnan ttaaatccnt gcctnggcga 720
agtccnttgn agggntaaan ggcccccтnn cggg 754

<210> 21
<211> 755
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(755)
<223> n = A,T,C or G

<400> 21
atcancccat gaccccnaac nngggaccnc tcanccggnc nnncnaccnc cggccnatca 60
nngtnagnnc actncnnttn natcacnccc cnccnactac gcccncnanc cnacgcncta 120
nncanatncc actganngcg cgangtngan ngagaaanct nataccanag ncaccanacn 180
ccagctgtcc nanaangcct nnnatacngg nnnatccaat ntgnancctc cnaagtattn 240
nncnncanat gattttcctn anccgattac ccntnccccc tanсccctcc cccccaacna 300
cgaaggcnct ggnccnaagg nngcgncncc ccgctagntc cccnncaagt cncncnccta 360
aactcanccn nattacncgc ttcntgagta tcactccccg aatctcaccc tactcaactc 420
aaaaanatcn gatacaaaat aatncaagcc tgnttatnac actntgactg ggtctctatt 480
ttagnggtcc ntnaancntc ctaatacttc cagtctncct tcnccaattt ccnaanggct 540
ctttcngaca gcatnttttg gttcccnntt gggttcttan ngaattgccc ttcntngaac 600
gggctcntct tttccttcgg ttancctggn ttcnnccggc cagttattat ttcccntttt 660
aaattcntnc cntttanttt tggcnttcna aacсcccggc cttgaaaacg gcсccctggt 720
aaaaggttgt tttganaaaa tttttgtttt gttcc 755

<210> 22
<211> 849
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(849)
<223> n = A,T,C or G

<400> 22
tttttttttt tttttangtg tngtcgtgca ggtagaggct tactacaant gtgaanacgt 60
acgctnggan taangcgacc cganttctag ganncnccct aaaatcanac tgtgaagatn 120

atcctgnnna cggaanggtc accggnngat nntgctaggg tgnccnctcc cannncnttn 180
cataactcng nggccctgcc caccaccttc ggcggcccng ngnccgggcc cgggtcattn 240
gnnttaaccn cactnngcna ncggtttccn nccccnncng acccnggcga tccggggtnc 300
tctgtcttcc cctgnagncn anaaantggg ccncggnccc ctttaccccc nnacaagcca 360
cngccntcta nccncngccc cccctccant nnggggggact gccnanngct ccgttnctng 420
nnaccccnnn gggtncctcg gttgtcgant cnaccgnang ccanggattc cnaaggaagg 480
tgcgttnttg gcccctaccc ttcgctncgg nncacccttc ccgacnanga nccgctcccg 540
cncnncgnng cctcncctcg caacacccgc nctcntcngt ncggnnnccc ccccacccgc 600
nccctcncnc ngncgnancn ctccncccnc gtctcannca ccaccccgcc ccgccaggcc 660
ntcanccacn ggnngacnng nagcncnntc gcnccgcgcn gcgncncccc cgccncngaa 720
ctncntcngg ccantnncgc tcaanccnna cnaaacgccg ctgcgcggcc cgnagcgncc 780
ncctccncga gtcctcccgn cttccnaccc angnnttccn cgaggacacn nnaccccgcc 840
nncangcgg 849

<210> 23
<211> 872
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(872)
<223> n = A,T,C or G

<400> 23
gcgcaaacta tacttcgctc gnactcgtgc gcctcgctnc tcttttcctc cgcaaccatg 60
tctgacnanc ccgattnggc ngatatcnan aagntcganc agtccaaact gantaacaca 120
cacacncnan aganaaatcc nctgccttcc anagtanacn attgaacnng agaaccangc 180
nggcgaatcg taatnaggcg tgcgccgcca atntgtcncc gtttattntn ccagcntcnc 240
ctnccnaccc tacntcttcn nagctgtcnn acccctngtn cgnacccccc naggtcggga 300
tcgggtttnn nntgaccgng cnnccccctcc ccccntccat nacganccnc ccgcaccacc 360
nanngcncgc nccccgnnct cttcgccncc ctgtcctntn cccctgtngc ctggcncngn 420
accgcattga ccctcgccnn ctncnngaaa ncgnanacgt ccgggttgnn annancgctg 480
tgggnnngcg tctgcnccgc gttccttccn ncnncttcca ccatcttcnt tacngggtct 540
ccncgccntc tcnnncacnc cctgggacgc tntcctntgc ccccccttnac tccccccctt 600
cgncgtgncc cgncccccacc ntcatttnca nacgntcttc acaannncct ggntnnctcc 660
cnancngncn gtcanccnag ggaagggngg ggnnccnntg nttgacgttg nggngangtc 720
cgaanantcc tcnccntcan cnctacccct cgggcgnnct ctcngttncc aacttancaa 780
ntctcccccg ngngcncntc tcagcctcnc ccncccccnct ctctgcantg tnctctgctc 840
tnaccnntac gantnttcgn cnccctcttt cc 872

<210> 24
<211> 815
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(815)
<223> n = A,T,C or G

<400> 24
gcatgcaagc ttgagtattc tatagngtca cctaaatanc ttggcntaat catggtcnta 60
nctgncttcc tgtgtcaaat gtatacnaan tanatatgaa tctnatntga caagannngta 120
tcntncatta gtaacaantg tnntgtccat cctgtcngan canattccca tnnattncgn 180
cgcattcncn gcncantatn taatngggaa ntcnnntnnn ncaccnncat ctatcntncc 240
gcncgctgac tggnagagat ggatnanttc tnntntgacc nacatgttca tcttggattn 300
aanancccgc cgcngnccac cggttngnng cnagccnntc ccaagacctc ctgtggaggt 360

aacctgcgtc aganncatca aacntgggaa acccgcnncc angtnnaagt ngnnncanan 420
gatcccgtcc aggnttnacc atcccttcnc agcgcccccc ttngtgcctt anagngnagc 480
gtgtccnanc cnctcaacat ganacgcgcc agnccanccg caattnggca caatgtcgnc 540
gaaccccccta gggggantna tncaaanccc caggattgtc cncncangaa atcccncanc 600
cccncccctac ccnnctttgg gacngtgacc aantcccgga gtnccagtcc ggccngnctc 660
ccccaccggt nnccntgggg gggtgaanct cngnntcanc cngncgaggn ntcgnaagga 720
accggncctn ggncgaannng ancnntcnga agngccncnt cgtataaccc cccctcncca 780
nccnacngnt agntcccccc cngggtncgg aangg 815

<210> 25
<211> 775
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(775)
<223> n = A,T,C or G

<400> 25
ccgagatgtc tcgctccgtg gccttagctg tgctcgcgct actctctctt tctggcctgg 60
aggctatcca gcgtactcca aagattcagg tttactcacg tcatccagca gagaatggaa 120
agtcaaattt cctgaattgc tatgtgtctg ggtttcatcc atccgacatt gaanttgact 180
tactgaagaa tgganagaga attgaaaaag tggagcattc agacttgtct ttcagcaagg 240
actggtcttt ctatctcntg tactacactg aattcacccc cactgaaaaa gatgagtatg 300
cctgccgtgt gaaccatgtg actttgtcac agcccaagat agttaagtgg gatcgagaca 360
tgtaagcagn cnncatggaa gtttgaagat gccgcatttg gattggatga attccaaatt 420
ctgcttgctt gcnttttaat antgatatgc ntatacaccc tacccttttat gnccccaaat 480
tgtaggggtt acatnantgt tcncntngga catgatcttc ctttataant ccnccnttcg 540
aattgcccgt cncccngttn ngaatgtttc cnnaaccacg gttggctccc ccaggtcncc 600
tcttacggaa gggcctgggc cnctttncaa ggttgggggа accnaaaatt tcncttntgc 660
ccncccncca cnntcttgng nncncanttt ggaacccttc cnattcccct tggcctcnna 720
nccttnncta anaaaacttn aaancgtngc naaannnttn acttcccccc ttacc 775

<210> 26
<211> 820
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(820)
<223> n = A,T,C or G

<400> 26
anattantac agtgtaatct tttcccagag gtgtgtanag ggaacggggc ctagaggcat 60
cccanagata ncttatanca acagtgcttt gaccaagagc tgctgggcac atttcctgca 120
gaaaaggtgg cggtccccat cactcctcct ctcccatagc catcccagag gggtgagtag 180
ccatcangcc ttcggtggga gggagtcang gaaacaacan accacagagc anacagacca 240
ntgatgacca tgggcgggag cgagcctctt ccctgnaccg gggtggcana nganagccta 300
nctgaggggt cacactataa acgttaacga ccnagatnan cacctgcttc aagtgcaccc 360
ttcctacctg acnaccagng accnnnaact gcngcctggg gacagcnctg ggancagcta 420
acnnagcact cacctgcccc cccatggccg tncgcntccc tggtcctgnc aagggaagct 480
ccctgttgga attncgggga naccaaggga nccccctcct ccanctgtga aggaaaaann 540
gatggaattt tnccccttccg gccnntcccc tcttccttta cacgcccccc nntactcntc 600
tccctctntt ntcctgncnc acttttnacc ccnnnatttc ccttnattga tcggannctn 660
ganattccac tnncgcctnc cntcnatcng naanacnaaa nactntctna cccnggggat 720
gggnncctcg ntcatcctct ctttttcnct accnccnntt ctttgcctct ccttngatca 780

tccaaccntc gntggccntn ccccccnnnn tcctttnccc 820

<210> 27
<211> 818
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(818)
<223> n = A,T,C or G

<400> 27
tctgggtgat ggcctcttcc tcctcaggga cctctgactg ctctgggcca aagaatctct 60
tgtttcttct ccgagcccca ggcagcggtg attcagccct gcccaacctg attctgatga 120
ctgcggatgc tgtgacggac ccaaggggca aatagggtcc cagggtccag ggaggggcgc 180
ctgctgagca cttccgcccc tcaccctgcc cagcccctgc catgagctct gggctgggtc 240
tccgcctcca gggttctgct cttccangca ngccancaag tggcgctggg ccacactggc 300
ttcttcctgc cccntccctg gctctgantc tctgtcttcc tgtcctgtgc angcnccttg 360
gatctcagtt tccctcnctc anngaactct gtttctgann tcttcantta actntgantt 420
tatnaccnan tggnctgtnc tgtcnnactt taatgggccn gaccggctaa tccctccctc 480
nctcccttcc anttcnnnna accngcttnc cntcntctcc ccntancccg ccngggaanc 540
ctcctttgcc ctnaccangg gccnnnaccg cccntnnctn gggggggcnng gtnnctncnc 600
ctgntnnccc cnctcncnnt tncctcgtcc cnncnncgcn nngcannttc ncngtcccnn 660
tnnctcttcn ngtntcgnaa ngntcncntn tnnnnngncn ngntnntncn tccctctcnc 720
cnnntgnang tnnttnnnnc ncngnncccc nnnncnnnnn nggnnntnnn tctncncngc 780
cccnnccccc ngnattaagg cctccnntct ccggccnc 818

<210> 28
<211> 731
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(731)
<223> n = A,T,C or G

<400> 28
aggaagggcg gagggatatt gtangggatt gagggatagg agnataangg gggaggtgtg 60
tcccaacatg anggtgnngt tctcttttga angagggttg ngtttttann ccnggtgggt 120
gattnaaccc cattgtatgg agnnaaaggn tttnagggat ttttcggctc ttatcagtat 180
ntanattcct gtnaatcgga aaatnatntt tcnncnggaa aatnttgctc ccatccgnaa 240
attnctcccg ggtagtgcat nttngggggn cngccangtt tcccaggctg ctanaatcgt 300
actaaagntt naagtgggan tncaaatgaa aacctnncac agagnatccn tacccgactg 360
tnnnttncct tcgccctntg actctgcnng agcccaatac ccnngngnat gtcncccngn 420
nnngcgncnc tgaaannnnc tcgnggctnn gancatcang gggtttcgca tcaaaagcnn 480
cgtttcncat naaggcactt tngcctcatc caaccnctng ccctcnncca tttngccgtc 540
nggttcncct acgctnntng cncctnnntn ganattttnc ccgcctnggg naancctcct 600
gnaatgggta gggncttntc ttttnaccnn gnggtntact aatcnnctnc acgcntnctt 660
tctcnacccc ccccctttt caatcccanc ggcnaatggg gtctccccnn cgangggggg 720
nnncccannc c 731

<210> 29
<211> 822
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(822)
<223> n = A,T,C or G

<400> 29
actagtccag tgtggtggaa ttccattgtg ttggggncnc ttctatgant antnttagat 60
cgctcanacc tcacancctc ccnacnangc ctataangaa nannaataga nctgtncnnt 120
atntntacnc tcatanncct cnnnacccac tccctcttaa cccntactgt gcctatngcn 180
tnnctantct ntgccgcctn cnanccaccn gtgggccnac cncnngnatt ctcnatctcc 240
tcnccatntn gcctananta ngtncatacc ctatacctac nccaatgcta nnnctaancn 300
tccatnantt annntaacta ccactgacnt ngactttcnc atnanctcct aatttgaatc 360
tactctgact cccacngcct annnattagc ancntccccc nacnatntct caaccaaatc 420
ntcaacaacc tatctanctg ttcnccaacc nttncctccg atccccnnac aacccccctc 480
ccaaataccc nccacctgac ncctaacccn caccatcccg gcaagccnan ggncatttan 540
ccactggaat cacnatngga naaaaaaaac ccnaactctc tancncnnat ctccctaana 600
aatnctcctn naatttactn ncantnccat caancccacn tgaaacnnaa cccctgtttt 660
tanatccctt ctttcgaaaa ccnacccttt annncccaac ctttngggcc ccccccnctnc 720
ccnaatgaag gncncccaat cnangaaacg nccntgaaaa ancnaggcna anannntccg 780
canatcctat cccttanttn ggggnccctt ncccngggcc cc 822

<210> 30
<211> 787
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(787)
<223> n = A,T,C or G

<400> 30
cggccgcctg ctctggcaca tgcctcctga atggcatcaa aagtgatgga ctgcccattg 60
ctagagaaga ccttctctcc tactgtcatt atggagccct gcagactgag ggctcccctt 120
gtctgcagga tttgatgtct gaagtcgtgg agtgtggctt ggagctcctc atctacatna 180
gctggaagcc ctggagggcc tctctcgcca gcctcccccc tctctccacg ctctccangg 240
acaccagggg ctccaggcag cccattattc ccagnangac atggtgtttc tccacgcgga 300
cccatggggc ctgnaaggcc agggtctcct ttgacaccat ctctcccgtc ctgcctggca 360
ggccgtggga tccactantt ctanaacggn cgccaccncg gtgggagctc cagcttttgt 420
tcccnttaat gaaggttaat tgcncgcttg gcgtaatcat nggtcanaac tntttcctgt 480
gtgaaattgt ttntcccctc ncnattccnc ncnacatacn aacccggaan cataaagtgt 540
taaagcctgg gggtngcctn nngaatnaac tnaactcaat taattgcgtt ggctcatggc 600
ccgctttccn ttcnggaaaa ctgtcntccc ctgcnttnnt gaatcggcca ccccccnggg 660
aaaagcggtt tgcnttttng ggggntcctt ccncttcccc cctcnctaan ccctncgcct 720
cggtcgttnc nggtngcggg gaangggnat nnnctcccnc naaggggggng agnnngntat 780
ccccaaa 787

<210> 31
<211> 799
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(799)
<223> n = A,T,C or G

<400> 31

```
ttttttttttt ttttttttggc gatgctactg tttaattgca ggaggtgggg gtgtgtgtac      60
catgtaccag ggctattaga agcaagaagg aaggagggag ggcagagcgc cctgctgagc      120
aacaaaggac tcctgcagcc ttctctgtct gtctcttggc gcaggcacat ggggaggcct      180
cccgcagggt gggggccacc agtccagggg tgggagcact acanggggtg ggagtgggtg      240
gtggctggtn cnaatggcct gncacanatc cctacgattc ttgacacctg gatttcacca      300
ggggaccttc tgttctccca nggnaacttc ntnnatctcn aaagaacaca actgtttctt      360
cngcanttct ggctgttcat ggaaagcaca ggtgtccnat ttnggctggg acttggtaca      420
tatggttccg gcccacctct ccentcnaan aagtaattca ccccccccn ccntctnttg      480
cctgggccct taantaccca caccggaact canttantta ttcatcttng gntgggcttg      540
ntnatcnccn cctgaangcg ccaagttgaa aggccacgcc gtncccnctc cccatagnan      600
nttttnncnt canctaatgc ccccccnggc aacnatccaa tccccccccn tgggggcccc      660
agcccanggc ccccgnctcg ggnnnccngn cncgnantcc ccaggntctc ccantcngnc      720
ccnnngcncc cccgcacgca gaacanaagg ntngagccnc cgcannnnnn nggtnncnac      780
ctcgcccccc ccnncgnng                                                   799
```

<210> 32
<211> 789
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(789)
<223> n = A,T,C or G

<400> 32

```
ttttttttttt ttttttttttt ttttttttttt ttttttttttt ttttttttttt ttttttttttt      60
ttttnccnag ggcaggttta ttgacaacct cncgggacac aancaggctg gggacaggac      120
ggcaacaggc tccggcggcg gcggcggcgg ccctacctgc ggtaccaaat ntgcagcctc      180
cgctcccgct tgatnttcct ctgcagctgc aggatgccnt aaaacagggc ctcggccntn      240
ggtgggcacc ctgggatttn aatttccacg ggcacaatgc ggtcgcancc cctcaccacc      300
nattaggaat agtggtntta cccnccnccg ttggcncact cccccntggaa accacttntc      360
gcggctccgg catctggtct taaaccttgc aaacnctggg gcccctcttt tggttantnt      420
nccngccaca atcatnactc agactggcnc gggctggccc caaaaaancn ccccaaaacc      480
ggnccatgtc ttnncggggt tgctgcnatn tncatcacct cccgggcnca ncaggncaac      540
ccaaaagttc ttgnggcccn caaaaaanct ccgggggggnc ccagtttcaa caaagtcatc      600
ccccttggcc cccaaatcct ccccccgntt nctgggtttg ggaacccacg cctctnnctt      660
tggnnggcaa gntggntccc ccttcgggcc cccggtgggc ccnnctctaa ngaaaacncc      720
ntcctnnnca ccatcccccc nngnnacgnc tancaangna tccctttttt tanaaacggg      780
ccccccncg                                                              789
```

<210> 33
<211> 793
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(793)
<223> n = A,T,C or G

<400> 33

```
gacagaacat gttggatggt ggagcacctt tctatacgac ttacaggaca gcagatgggg      60
aattcatggc tgttggagca atanaacccc agttctacga gctgctgatc aaaggacttg      120
gactaaagtc tgatgaactt cccaatcaga tgagcatgga tgattggcca gaaatgaana      180
agaagtttgc agatgtattt gcaaagaaga cgaaggcaga gtggtgtcaa atctttgacg      240
gcacagatgc ctgtgtgact ccggttctga cttttgagga ggttgttcat catgatcaca      300
acaangaacg gggctcgttt atcaccantg aggagcagga cgtgagcccc cgccctgcac      360
```

ctctgctgtt aaacacccca gccatccctt ctttcaaaag ggatccacta cttctagagc 420
ggncgccacc gcggtggagc tccagctttt gttcccttta gtgagggtta attgcgcgct 480
tggcgtaatc atggtcatan ctgtttcctg tgtgaaattg ttatccgctc acaattccac 540
acaacatacg anccggaagc atnaaatttt aaagcctggn ggtngcctaa tgantgaact 600
nactcacatt aattggcttt gcgctcactg cccgctttcc agtccggaaa acctgtcctt 660
gccagctgcc nttaatgaat cnggccaccc cccggggaaa aggcngtttg cttnttgggg 720
cgcncttccc gctttctcgc ttcctgaant ccttcccccc ggtctttcgg cttgcggcna 780
acggtatcna cct 793

<210> 34
<211> 756
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(756)
<223> n = A,T,C or G

<400> 34
gccgcgaccg gcatgtacga gcaactcaag ggcgagtgga accgtaaaag ccccaatctt 60
ancaagtgcg gggaanagct gggtcgactc aagctagttc ttctggagct caacttcttg 120
ccaaccacag ggaccaagct gaccaaacag cagctaattc tggcccgtga catactggag 180
atcggggccc aatggagcat cctacgcaan gacatcccct ccttcgagcg ctacatggcc 240
cagctcaaat gctactactt tgattacaan gagcagctcc ccgagtcagc ctatatgcac 300
cagctcttgg gcctcaacct cctcttcctg ctgtcccaga accgggtggc tgantnccac 360
acgganttgg ancggctgcc tgcccaanga catacanacc aatgtctaca tcnaccacca 420
gtgtcctgga gcaatactga tgganggcag ctaccncaaa gtnttcctgg ccnagggtaa 480
catcccccgc cgagagctac accttcttca ttgacatcct gctcgacact atcagggatg 540
aaaatcgcng ggttgctcca gaaaggctnc aanaanatcc ttttcnctga aggccccccgg 600
atncnctagt nctagaatcg gcccgccatc gcggtgganc ctccaacctt tcgttnccct 660
ttactgaggg ttnattgccg cccttggcgt tatcatggtc acnccngttn cctgtgttga 720
aattnttaac cccccacaat tccacgccna cattng 756

<210> 35
<211> 834
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(834)
<223> n = A,T,C or G

<400> 35
ggggatctct anatcnacct gnatgcatgg ttgtcggtgt ggtcgctgtc gatgaanatg 60
aacaggatct tgcccttgaa gctctcggct gctgtnttta agttgctcag tctgccgtca 120
tagtcagaca cnctcttggg caaaaaacan caggatntga gtcttgattt cacctccaat 180
aatcttcngg gctgtctgct cggtgaactc gatgacnang ggcagctggt tgtgtntgat 240
aaantccanc angttctcct tggtgacctc cccttcaaag ttgttccggc cttcatcaaa 300
cttctnnaan angannancc cancttttgtc gagctggnat ttgganaaca cgtcactgtt 360
ggaaactgat cccaaatggt atgtcatcca tcgcctctgc tgcctgcaaa aaacttgctt 420
ggcncaaatc cgactccccn tccttgaaag aagccnatca cacccccctc cctggactcc 480
nncaangact ctnccgctnc cccntccnng cagggttggt ggcannccgg gccccntgcgc 540
ttcttcagcc agttcacnat nttcatcagc ccctctgcca gctgttntat tccttggggg 600
ggaanccgtc tctcccttcc tgaannaact ttgaccgtng gaatagccgc gcntcnccnt 660
acntnctggg ccgggttcaa antccctccn ttgncnntcn cctcgggcca ttctggattt 720
nccnaacttt ttccttcccc cnccccncgg ngtttggntt tttcatnggg ccccaactct 780

gctnttggcc antcccctgg gggcntntan cnccccctnt ggtcccntng ggcc 834

<210> 36
<211> 814
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(814)
<223> n = A,T,C or G

<400> 36
cggncgcttt ccngccgcgc cccgtttcca tgacnaaggc tcccttcang ttaaatacnn 60
cctagnaaac attaatgggt tgctctacta atacatcata cnaaccagta agcctgccca 120
naacgccaac tcaggccatt cctaccaaag gaagaaaggc tggtctctcc acccccctgta 180
ggaaaggcct gccttgtaag acaccacaat ncggctgaat ctnaagtctt gtgttttact 240
aatggaaaaa aaaaataaac aanaggtttt gttctcatgg ctgcccaccg cagcctggca 300
ctaaaacanc ccagcgctca cttctgcttg ganaaatatt ctttgctctt ttggacatca 360
ggcttgatgg tatcactgcc acntttccac ccagctgggc nccctttccc catntttgtc 420
antganctgg aaggcctgaa ncttagtctc caaaagtctc ngcccacaag accggccacc 480
aggggangtc ntttncagtg gatctgccaa anantacccn tatcatcnnt gaataaaaag 540
gcccctgaac ganatgcttc cancancctt taagacccat aatcctngaa ccatggtgcc 600
cttccggtct gatccnaaag gaatgttcct gggtcccant ccctcctttg ttncttacgt 660
tgtnttggac ccntgctngn atnacccaan tganatcccc ngaagcaccc tnccccctggc 720
atttganttt cntaaattct ctgccctacn nctgaaagca cnattccctn ggcnccnaan 780
ggngaactca agaaggtctn ngaaaaacca cncn 814

<210> 37
<211> 760
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(760)
<223> n = A,T,C or G

<400> 37
gcatgctgct cttcctcaaa gttgttcttg ttgccataac aaccaccata ggtaaagcgg 60
gcgcagtgtt cgctgaaggg gttgtagtac cagcgcggga tgctctcctt gcagagtcct 120
gtgtctggca ggtccacgca atgccctttg tcactgggga aatggatgcg ctggagctcg 180
tcnaanccac tcgtgtattt ttcacangca gcctcctccg aagcntccgg gcagttgggg 240
gtgtcgtcac actccactaa actgtcgatn cancagccca ttgctgcagc ggaactgggt 300
gggctgacag gtgccagaac acactggatn ggcctttcca tggaagggcc tgggggaaat 360
cnnctnancc caaactgcct ctcaaaggcc accttgcaca ccccgacagg ctagaaatgc 420
actcttcttc ccaaaggtag ttgttcttgt tgcccaagca ncctccanca aaccaaaanc 480
ttgcaaaatc tgctccgtgg gggtcatnnn taccanggtt ggggaaanaa acccggcngn 540
ganccncctt gtttgaatgc naaggnaata atcctcctgt cttgcttggg tggaanagca 600
caattgaact gttaacnttg ggccgngttc cnctngggtg gtctgaaact aatcaccgtc 660
actggaaaaa ggtangtgcc ttccttgaat tcccaaantt cccctngntt tgggtnnttt 720
ctcctctncc ctaaaaatcg tnttcccccc ccntanggcg 760

<210> 38
<211> 724
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(724)
<223> n = A,T,C or G

<400> 38
ttttttttttt ttttttttttt ttttttttttt ttttaaaaa ccccctccat tgaatgaaaa 60
cttccnaaat tgtccaaccc cctcnnccaa atnnccattt ccgggggggg gttccaaacc 120
caaattaatt ttgganttta aattaaatnt tnattngggg aanaanccaa atgtnaagaa 180
aatttaaccc attatnaact taaatncctn gaaaccccntg gnttccaaaa atttttaacc 240
cttaaatccc tccgaaattg ntaanggaaa accaaattcn cctaaggctn tttgaaggtt 300
ngatttaaac ccccttnant tnttttnacc cnngnctnaa ntatttngnt tccggtgttt 360
tcctnttaan cntnggtaac tcccgntaat gaannnccct aanccaatta aaccgaattt 420
tttttgaatt ggaaattccn ngggaattna ccggggtttt tcccntttgg gggccatncc 480
cccnctttcg gggtttgggn ntaggttgaa tttttnnang ncccaaaaaa nccccccaana 540
aaaaaactcc caagnnttaa ttngaatntc ccccttccca ggccttttgg gaaaggnggg 600
tttntggggg ccngggantt cnttcccccn ttnccncccc ccccccnggt aaanggttat 660
ngnntttggt ttttgggccc cttnanggac cttccggatn gaaattaaat ccccgggncg 720
gccg 724

<210> 39
<211> 751
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(751)
<223> n = A,T,C or G

<400> 39
ttttttttttt tttttctttg ctcacattta atttttattt tgatttttttt taatgctgca 60
caacacaata tttatttcat ttgtttcttt tatttcattt tatttgtttg ctgctgctgt 120
tttatttatt tttactgaaa gtgagaggga acttttgtgg ccttttttcc tttttctgta 180
ggccgcctta agctttctaa atttggaaca tctaagcaag ctgaanggaa aaggggggttt 240
cgcaaaatca ctcgggggaa nggaaaggtt gctttgttaa tcatgcccta tggtgggtga 300
ttaactgctt gtacaattac ntttcacttt taattaattg tgctnaangc tttaattana 360
cttgggggtt ccctccccan accaaccccn ctgacaaaaa gtgccngccc tcaaatnatg 420
tcccggcnnt cnttgaaaca cacngcngaa ngttctcatt ntccccncnc caggtnaaaa 480
tgaagggtta ccatntttaa cnccacctcc acntggcnnn gcctgaatcc tcnaaaancn 540
ccctcaancn aattnctnng ccccggtcnc gcntnngtcc cncccgggct ccgggaantn 600
caccccnga anncnntnnc naacnaaatt ccgaaaatat tcccnntcnc tcaattcccc 660
cnnagactnt cctcnncnan cncaattttc ttttnntcac gaacncgnnc cnnaaaatgn 720
nnnncncctc cnctngtccn naatcnccan c 751

<210> 40
<211> 753
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(753)
<223> n = A,T,C or G

<400> 40
gtggtatttt ctgtaagatc aggtgttcct ccctcgtagg tttagaggaa acaccctcat 60
agatgaaaac ccccccgaga cagcagcact gcaactgcca agcagccggg gtaggagggg 120

cgccctatgc acagctgggc ccttgagaca gcagggcttc gatgtcaggc tcgatgtcaa 180
tggtctggaa gcggcggctg tacctgcgta ggggcacacc gtcagggccc accaggaact 240
tctcaaagtt ccaggcaacn tcgttgcgac acaccggaga ccaggtgatn agcttggggt 300
cggtcataan cgcggtggcg tcgtcgctgg gagctggcag ggcctcccgc aggaaggcna 360
ataaaaggtg cgcccccgca ccgttcanct cgcacttctc naanaccatg angttgggct 420
cnaacccacc accannccgg acttccttga nggaattccc aaatctcttc gntcttgggc 480
ttctnctgat gccctanctg gttgcccngn atgccaanca nccccaancc ccggggtcct 540
aaancacccn cctcctcntt tcatctgggt tnttntcccc ggaccntggt tcctctcaag 600
ggancccata tctcnaccan tactcaccnt nccccccccnt gnnacccanc cttctanngn 660
ttcccncccg ncctctggcc cntcaaanan gcttncacna cctgggtctg ccttcccccc 720
tnccctatct gnaccccncn tttgtctcan tnt 753

<210> 41
<211> 341
<212> DNA
<213> Homo sapien

<400> 41
actatatcca tcacaacaga catgcttcat cccatagact tcttgacata gcttcaaatg 60
agtgaaccca tccttgattt atatacatat atgttctcag tattttggga gcctttccac 120
ttctttaaac cttgttcatt atgaacactg aaaataggaa tttgtgaaga gttaaaaagt 180
tatagcttgt ttacgtagta agtttttgaa gtctacattc aatccagaca cttagttgag 240
tgttaaactg tgatttttaa aaaatatcat ttgagaatat tctttcagag gtattttcat 300
ttttactttt tgattaattg tgttttatat attagggtag t 341

<210> 42
<211> 101
<212> DNA
<213> Homo sapien

<400> 42
acttactgaa tttagttctg tgctcttcct tatttagtgt tgtatcataa atactttgat 60
gtttcaaaca ttctaaataa ataatttttca gtggcttcat a 101

<210> 43
<211> 305
<212> DNA
<213> Homo sapien

<400> 43
acatctttgt tacagtctaa gatgtgttct taaatcacca ttccttcctg gtcctcaccc 60
tccagggtgg tctcacactg taattagagc tattgaggag tctttacagc aaattaagat 120
tcagatgcct tgctaagtct agagttctag agttatgttt cagaaagtct aagaaaccca 180
cctcttgaga ggtcagtaaa gaggacttaa tatttcatat ctacaaaatg accacaggat 240
tggatacaga acgagagtta tcctggataa ctcagagctg agtacctgcc cgggggccgc 300
tcgaa 305

<210> 44
<211> 852
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(852)
<223> n = A,T,C or G

<400> 44

acataaatat cagagaaaag tagtctttga aatatttacg tccaggagtt ctttgtttct 60
gattatttgg tgtgtgtttt ggtttgtgtc caaagtattg gcagcttcag ttttcatttt 120
ctctccatcc tcgggcattc ttcccaaatt tatataccag tcttcgtcca tccacacgct 180
ccagaatttc tcttttgtag taatatctca tagctcggct gagcttttca taggtcatgc 240
tgctgttgtt cttcttttta ccccatagct gagccactgc ctctgatttc aagaacctga 300
agacgccctc agatcggtct tcccatttta ttaatcctgg gttcttgtct gggttcaaga 360
ggatgtcgcg gatgaattcc cataagtgag tccctctcgg gttgtgcttt ttggtgtggc 420
acttggcagg ggggtcttgc tccttttca tatcaggtga ctctgcaaca ggaaggtgac 480
tggtggttgt catggagatc tgagcccggc agaaagtttt gctgtccaac aaatctactg 540
tgctaccata gttggtgtca tataaatagt tctngtcttt ccaggtgttc atgatggaag 600
gctcagtttg ttcagtcttg acaatgacat tgtgtgtgga ctggaacagg tcactactgc 660
actggccgtt ccacttcaga tgctgcaagt tgctgtagag gagntgcccc gccgtccctg 720
ccgcccgggt gaactcctgc aaactcatgc tgcaaaggtg ctcgccgttg atgtcgaact 780
cntggaaagg gatacaattg gcatccagct ggttggtgtc caggaggtga tggagccact 840
cccacacctg gt 852

<210> 45
<211> 234
<212> DNA
<213> Homo sapien

<400> 45
acaacagacc cttgctcgct aacgacctca tgctcatcaa gttggacgaa tccgtgtccg 60
agtctgacac catccggagc atcagcattg cttcgcagtg ccctaccgcg gggaactctt 120
gcctcgtttc tggctggggt ctgctggcga acggcagaat gcctaccgtg ctgcagtgcg 180
tgaacgtgtc ggtggtgtct gaggaggtct gcagtaagct ctatgacccg ctgt 234

<210> 46
<211> 590
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(590)
<223> n = A,T,C or G

<400> 46
acttttttatt taaatgttta taaggcagat ctatgagaat gatagaaaac atggtgtgta 60
atttgatagc aatattttgg agattacaga gttttagtaa ttaccaatta cacagttaaa 120
aagaagataa tatattccaa gcanatacaa aatatctaat gaaagatcaa ggcaggaaaa 180
tgantataac taattgacaa tggaaaatca attttaatgt gaattgcaca ttatccttta 240
aaagctttca aaanaaanaa ttattgcagt ctanttaatt caaacagtgt taaatggtat 300
caggataaan aactgaaggg canaaagaat taattttcac ttcatgtaac ncacccanat 360
ttacaatggc ttaaatgcan ggaaaaagca gtggaagtag ggaagtantc aaggtctttc 420
tggtctctaa tctgccttac tctttgggtg tggctttgat cctctggaga cagctgccag 480
ggctcctgtt atatccacaa tcccagcagc aagatgaagg gatgaaaaag gacacatgct 540
gccttccttt gaggagactt catctcactg gccaacactc agtcacatgt 590

<210> 47
<211> 774
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(774)
<223> n = A,T,C or G

<400> 47
acaaggggc ataatgaagg agtggggana gattttaaag aaggaaaaaa aacgaggccc 60
tgaacagaat tttcctgnac aacggggctt caaaataatt ttcttgggga ggttcaagac 120
gcttcactgc ttgaaactta aatggatgtg ggacanaatt ttctgtaatg accctgaggg 180
cattacagac gggactctgg gaggaaggat aaacagaaag gggacaaagg ctaatcccaa 240
aacatcaaag aaaggaaggt ggcgtcatac ctcccagcct acacagttct ccagggctct 300
cctcatccct ggaggacgac agtggaggaa caactgacca tgtccccagg ctcctgtgtg 360
ctggctcctg gtcttcagcc cccagctctg gaagcccacc ctctgctgat cctgcgtggc 420
ccacactcct tgaacacaca tccccaggtt atattcctgg acatggctga acctcctatt 480
cctacttccg agatgccttg ctccctgcag cctgtcaaaa tcccactcac cctccaaacc 540
acggcatggg aagcctttct gacttgcctg attactccag catcttggaa caatccctga 600
ttccccactc cttagaggca agatagggtg gttaagagta gggctggacc acttggagcc 660
aggctgctgg cttcaaattn tggctcattt acgagctatg ggaccttggg caagtnatct 720
tcacttctat gggcntcatt ttgttctacc tgcaaaatgg gggataataa tagt 774

<210> 48
<211> 124
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(124)
<223> n = A,T,C or G

<400> 48
canaaattga aattttataa aaaggcattt ttctcttata tccataaaat gatataattt 60
ttgcaantat anaaatgtgt cataaattat aatgttcctt aattacagct caacgcaact 120
tggt 124

<210> 49
<211> 147
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(147)
<223> n = A,T,C or G

<400> 49
gccgatgcta ctattttatt gcaggaggtg ggggtgtttt tattattctc tcaacagctt 60
tgtggctaca ggtggtgtct gactgcatna aaaantttt tacgggtgat tgcaaaaatt 120
ttagggcacc catatcccaa gcantgt 147

<210> 50
<211> 107
<212> DNA
<213> Homo sapien

<400> 50
acattaaatt aataaaagga ctgttggggt tctgctaaaa cacatggctt gatatattgc 60
atggtttgag gttaggagga gttaggcata tgttttggga gaggggt 107

<210> 51
<211> 204
<212> DNA

<213> Homo sapien

<400> 51
gtcctaggaa gtctagggga cacacgactc tggggtcacg gggccgacac acttgcacgg 60
cgggaaggaa aggcagagaa gtgacaccgt caggggggaaa tgacagaaag gaaaatcaag 120
gccttgcaag gtcagaaagg ggactcaggg cttccaccac agccctgccc cacttggcca 180
cctccctttt gggaccagca atgt 204

<210> 52
<211> 491
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(491)
<223> n = A,T,C or G

<400> 52
acaaagataa catttatctt ataacaaaaa tttgatagtt ttaaaggtta gtattgtgta 60
gggtattttc caaaagacta aagagataac tcaggtaaaa agttagaaat gtataaaaca 120
ccatcagaca ggtttttaaa aaacaacata ttacaaaatt agacaatcat ccttaaaaaa 180
aaaacttctt gtatcaattt cttttgttca aaatgactga cttaantatt tttaaatatt 240
tcanaaacac ttcctcaaaa attttcaana tggtagcttt canatgtncc ctcagtccca 300
atgttgctca gataaataaa tctcgtgaga acttaccacc caccacaagc tttctggggc 360
atgcaacagt gtcttttctt tnctttttct ttttttttt ttacaggcac agaaactcat 420
caatttatt tggataacaa agggtctcca aattatattg aaaaataaat ccaagttaat 480
atcactcttg t 491

<210> 53
<211> 484
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(484)
<223> n = A,T,C or G

<400> 53
acataattta gcagggctaa ttaccataag atgctattta ttaanaggtn tatgatctga 60
gtattaacag ttgctgaagt ttggtatttt tatgcagcat tttctttttg ctttgataac 120
actacagaac ccttaaggac actgaaaatt agtaagtaaa gttcagaaac attagctgct 180
caatcaaatc tctacataac actatagtaa ttaaaacgtt aaaaaaaagt gttgaaatct 240
gcactagtat anaccgctcc tgtcaggata anactgcttt ggaacagaaa gggaaaaanc 300
agctttgant ttctttgtgc tgatangagg aaaggctgaa ttaccttgtt gcctctccct 360
aatgattggc aggtcnggta aatnccaaaa catattccaa ctcaacactt cttttccncg 420
tancttgant ctgtgtattc caggancagg cggatggaat gggccagccc ncggatgttc 480
cant 484

<210> 54
<211> 151
<212> DNA
<213> Homo sapien

<400> 54
actaaacctc gtgcttgtga actccataca gaaaacggtg ccatccctga acacggctgg 60
ccactgggta tactgctgac aaccgcaaca acaaaaacac aaatccttgg cactggctag 120

tctatgtcct ctcaagtgcc tttttgtttg t 151

<210> 55
<211> 91
<212> DNA
<213> Homo sapien

<400> 55
acctggcttg tctccgggtg gttcccggcg ccccccacgg tccccagaac ggacactttc 60
gccctccagt ggatactcga gccaaagtgg t 91

<210> 56
<211> 133
<212> DNA
<213> Homo sapien

<400> 56
ggcggatgtg cgttggttat atacaaatat gtcattttat gtaagggact tgagtatact 60
tggatttttg gtatctgtgg gttgggggga cggtccagga accaataccc catggatacc 120
aagggacaac tgt 133

<210> 57
<211> 147
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(147)
<223> n = A,T,C or G

<400> 57
actctggaga acctgagccg ctgctccgcc tctgggatga ggtgatgcan gcngtggcgc 60
gactgggagc tgagcccttc cctttgcgcc tgcctcagag gattgttgcc gacntgcana 120
tctcantggg ctggatncat gcagggt 147

<210> 58
<211> 198
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(198)
<223> n = A,T,C or G

<400> 58
acagggatat aggtttnaag ttattgtnat tgtaaaatac attgaatttt ctgtatactc 60
tgattacata catttatcct ttaaaaaaga tgtaaatctt aatttttatg ccatctatta 120
atttaccaat gagttacctt gtaaatgaga agtcatgata gcactgaatt ttaactagtt 180
ttgacttcta agtttggt 198

<210> 59
<211> 330
<212> DNA
<213> Homo sapien

<400> 59

acaacaaatg ggttgtgagg aagtcttatc agcaaaactg gtgatggcta ctgaaaagat 60
ccattgaaaa ttatcattaa tgattttaaa tgacaagtta tcaaaaactc actcaatttt 120
cacctgtgct agcttgctaa aatgggagtt aactctagag caaatatagt atcttctgaa 180
tacagtcaat aaatgacaaa gccagggcct acaggtggtt tccagacttt ccagacccag 240
cagaaggaat ctattttatc acatggatct ccgtctgtgc tcaaaatacc taatgatatt 300
tttcgtcttt attggacttc tttgaagagt 330

<210> 60
<211> 175
<212> DNA
<213> Homo sapien

<400> 60
accgtgggtg ccttctacat tcctgacggc tccttcacca acatctggtt ctacttcggc 60
gtcgtgggct ccttcctctt catcctcatc cagctggtgc tgctcatcga ctttgcgcac 120
tcctggaacc agcggtggct gggcaaggcc gaggagtgcg attcccgtgc ctggt 175

<210> 61
<211> 154
<212> DNA
<213> Homo sapien

<400> 61
accccacttt tcctcctgtg agcagtctgg acttctcact gctacatgat gagggtgagt 60
ggttgttgct cttcaacagt atcctcccct ttccggatct gctgagccgg acagcagtgc 120
tggactgcac agccccgggg ctccacattg ctgt 154

<210> 62
<211> 30
<212> DNA
<213> Homo sapien

<400> 62
cgctcgagcc ctatagtgag tcgtattaga 30

<210> 63
<211> 89
<212> DNA
<213> Homo sapien

<400> 63
acaagtcatt tcagcaccct ttgctcttca aaactgacca tcttttatat ttaatgcttc 60
ctgtatgaat aaaaatggtt atgtcaagt 89

<210> 64
<211> 97
<212> DNA
<213> Homo sapien

<400> 64
accggagtaa ctgagtcggg acgctgaatc tgaatccacc aataaataaa ggttctgcag 60
aatcagtgca tccaggattg gtccttggat ctggggt 97

<210> 65
<211> 377
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(377)
<223> n = A,T,C or G

<400> 65
acaacaanaa ntcccttctt taggccactg atggaaacct ggaaccccct tttgatggca 60
gcatggcgtc ctaggccttg acacagcggc tggggtttgg gctntcccaa accgcacacc 120
ccaaccctgg tctacccaca nttctggcta tgggctgtct ctgccactga acatcagggt 180
tcggtcataa natgaaatcc caanggggac agaggtcagt agaggaagct caatgagaaa 240
ggtgctgttt gctcagccag aaaacagctg cctggcattc gccgctgaac tatgaacccg 300
tgggggtgaa ctacccccan gaggaatcat gcctgggcga tgcaanggtg ccaacaggag 360
gggcgggagg agcatgt 377

<210> 66
<211> 305
<212> DNA
<213> Homo sapien

<400> 66
acgcctttcc ctcagaattc agggaagaga ctgtcgcctg ccttcctccg ttgttgcgtg 60
agaacccgtg tgcccttcc caccatatcc accctcgctc catctttgaa ctcaaacacg 120
aggaactaac tgcaccctgg tcctctcccc agtccccagt tcaccctcca tccctcacct 180
tcctccactc taagggatat caacactgcc cagcacaggg gccctgaatt tatgtggttt 240
ttatatattt tttaataaga tgcactttat gtcatttttt aataaagtct gaagaattac 300
tgttt 305

<210> 67
<211> 385
<212> DNA
<213> Homo sapien

<400> 67
actacacaca ctccacttgc ccttgtgaga cactttgtcc cagcacttta ggaatgctga 60
ggtcggacca gccacatctc atgtgcaaga ttgcccagca gacatcaggt ctgagagttc 120
ccctttaaa aaaggggact tgcttaaaaa agaagtctag ccacgattgt gtagagcagc 180
tgtgctgtgc tggagattca cttttgagag agttctcctc tgagacctga tctttagagg 240
ctgggcagtc ttgcacatga gatggggctg gtctgatctc agcactcctt agtctgcttg 300
cctctcccag ggccccagcc tggccacacc tgcttacagg gcactctcag atgcccatac 360
catagtttct gtgctagtgg accgt 385

<210> 68
<211> 73
<212> DNA
<213> Homo sapien

<400> 68
acttaaccag atatattttt accccagatg gggatattct ttgtaaaaaa tgaaaataaa 60
gtttttttaa tgg 73

<210> 69
<211> 536
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(536)

<223> n = A,T,C or G

<400> 69
actagtccag tgtggtggaa ttccattgtg ttgggggctc tcaccctcct ctcctgcagc 60
tccagctttg tgctctgcct ctgaggagac catggcccag catctgagta ccctgctgct 120
cctgctggcc accctagctg tggccctggc ctggagcccc aaggaggagg ataggataat 180
cccgggtggc atctataacg cagacctcaa tgatgagtgg gtacagcgtg cccttcactt 240
cgccatcagc gagtataaca aggccaccaa agatgactac tacagacgtc cgctgcgggt 300
actaagagcc aggcaacaga ccgttggggg ggtgaattac ttcttcgacg tagaggtggg 360
ccgaaccata tgtaccaagt cccagcccaa cttggacacc tgtgccttcc atgaacagcc 420
agaactgcag aagaaacagt tgtgctcttt cgagatctac gaagttccct ggggagaaca 480
gaangtccct gggtgaaatc caggtgtcaa gaaatcctan ggatctgttg ccaggc 536

<210> 70
<211> 477
<212> DNA
<213> Homo sapien

<400> 70
atgaccccta acaggggccc tctcagccct cctaatgacc tccggcctag ccatgtgatt 60
tcacttccac tccataacgc tcctcatact aggcctacta accaacacac taaccatata 120
ccaatgatgg cgcgatgtaa cacgagaaag cacataccaa ggccaccaca caccacctgt 180
ccaaaaaggc cttcgatacg ggataatcct atttattacc tcagaagttt ttttcttcgc 240
agggattttt ctgagccttt taccactcca gcctagcccc tacccccaa ctaggagggc 300
actggccccc aacaggcatc accccgctaa atcccctaga agtcccactc ctaaacacat 360
ccgtattact cgcatcagga gtatcaatca cctgagctca ccatagtcta atagaaaaca 420
accgaaacca aattattcaa agcactgctt attacaattt tactgggtct ctatttt 477

<210> 71
<211> 533
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(533)
<223> n = A,T,C or G

<400> 71
agagctatag gtacagtgtg atctcagctt tgcaaacaca ttttctacat agatagtact 60
aggtattaat agatatgtaa agaaagaaat cacaccatta ataatggtaa gattggttta 120
tgtgatttta gtggtatttt tggcaccctt atatatgttt tccaaacttt cagcagtgat 180
attatttcca taacttaaaa agtgagtttg aaaaagaaaa tctccagcaa gcatctcatt 240
taaataaagg tttgtcatct ttaaaaatac agcaatatgt gacttttaa aaaagctgtc 300
aaataggtgt gaccctacta ataattatta gaaatacatt taaaaacatc gagtacctca 360
agtcagtttg ccttgaaaaa tatcaaatat aactcttaga gaaatgtaca taaaagaatg 420
cttcgtaatt ttggagtang aggttccctc ctcaattttg tatttttaaa aagtacatgg 480
taaaaaaaaa aattcacaac agtatataag gctgtaaaat gaagaattct gcc 533

<210> 72
<211> 511
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(511)
<223> n = A,T,C or G

<400> 72
tattacggaa aaacacacca cataattcaa ctancaaaga anactgcttc agggcgtgta 60
aaatgaaagg cttccaggca gttatctgat taaagaacac taaaagaggg acaaggctaa 120
aagccgcagg atgtctacac tatancaggc gctatttggg ttggctggag gagctgtgga 180
aaacatggan agattggtgc tgganatcgc cgtggctatt cctcattgtt attacanagt 240
gaggttctct gtgtgcccac tggtttgaaa accgttctnc aataatgata gaatagtaca 300
cacatgagaa ctgaaatggc ccaaacccag aaagaaagcc caactagatc ctcagaanac 360
gcttctaggg acaataaccg atgaagaaaa gatggcctcc ttgtgccccc gtctgttatg 420
atttctctcc attgcagcna naaacccgtt cttctaagca aacncaggtg atgatggcna 480
aaatacaccc cctcttgaag naccnggagg a 511

<210> 73
<211> 499
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(499)
<223> n = A,T,C or G

<400> 73
cagtgccagc actggtgcca gtaccagtac caataacagt gccagtgcca gtgccagcac 60
cagtggtggc ttcagtgctg gtgccagcct gaccgccact ctcacatttg ggctcttcgc 120
tggccttggt ggagctggtg ccagcaccag tggcagctct ggtgcctgtg gtttctccta 180
caagtgagat tttagatatt gttaatcctg ccagtctttc tcttcaagcc agggtgcatc 240
ctcagaaacc tactcaacac agcactctag gcagccacta tcaatcaatt gaagttgaca 300
ctctgcatta aatctatttg ccatttctga aaaaaaaaaa aaaaaaaggg cggccgctcg 360
antctagagg gcccgtttaa acccgctgat cagcctcgac tgtgccttct anttgccagc 420
catctgttgt ttgcccctcc cccgntgcct tccttgaccc tggaaagtgc cactcccact 480
gtcctttcct aantaaaat 499

<210> 74
<211> 537
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(537)
<223> n = A,T,C or G

<400> 74
tttcatagga gaacacactg aggagatact tgaagaattt ggattcagcc gcgaagagat 60
ttatcagctt aactcagata aaatcattga aagtaataag gtaaaagcta gtctctaact 120
tccaggccca cggctcaagt gaatttgaat actgcattta cagtgtagag taacacataa 180
cattgtatgc atggaaacat ggaggaacag tattacagtg tcctaccact ctaatcaaga 240
aaagaattac agactctgat tctacagtga tgattgaatt ctaaaaatgg taatcattag 300
ggcttttgat ttataanact ttgggtactt atactaaatt atggtagtta tactgccttc 360
cagtttgctt gatatatttg ttgatattaa gattcttgac ttatattttg aatgggttct 420
actgaaaaan gaatgatata ttcttgaaga catcgatata catttattta cactcttgat 480
tctacaatgt agaaaatgaa ggaaatgccc caaattgtat ggtgataaaa gtcccgt 537

<210> 75
<211> 467
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(467)
<223> n = A,T,C or G

<400> 75
caaanacaat tgttcaaaag atgcaaatga tacactactg ctgcagctca caaacacctc 60
tgcatattac acgtacctcc tcctgctcct caagtagtgt ggtctatttt gccatcatca 120
cctgctgtct gcttagaaga acggctttct gctgcaangg agagaaatca taacagacgg 180
tggcacaagg aggccatctt ttcctcatcg gttattgtcc ctagaagcgt cttctgagga 240
tctagttggg ctttctttct gggtttgggc catttcantt ctcatgtgtg tactattcta 300
tcattattgt ataacggttt tcaaaccngt gggcacncag agaacctcac tctgtaataa 360
caatgaggaa tagccacggt gatctccagc accaaatctc tccatgttnt tccagagctc 420
ctccagccaa cccaaatagc cgctgctatn gtgtagaaca tccctgn 467

<210> 76
<211> 400
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(400)
<223> n = A,T,C or G

<400> 76
aagctgacag cattcgggcc gagatgtctc gctccgtggc cttagctgtg ctcgcgctac 60
tctctctttc tggcctggag gctatccagc gtactccaaa gattcaggtt tactcacgtc 120
atccagcaga gaatggaaag tcaaatttcc tgaattgcta tgtgtctggg tttcatccat 180
ccgacattga agttgactta ctgaagaatg gagagagaat tgaaaaagtg gagcattcag 240
acttgtcttt cagcaaggac tggtctttct atctcttgta ctacactgaa ttcaccccca 300
ctgaaaaaga tgagtatgcc tgccgtgtga accatgtgac tttgtcacag cccaagatng 360
ttnagtggga tcganacatg taagcagcan catgggaggt 400

<210> 77
<211> 248
<212> DNA
<213> Homo sapien

<400> 77
ctggagtgcc ttggtgtttc aagcccctgc aggaagcaga atgcaccttc tgaggcacct 60
ccagctgccc cggcgggga tgcgaggctc ggagcaccct tgcccggctg tgattgctgc 120
caggcactgt tcatctcagc ttttctgtcc ctttgctccc ggcaagcgct tctgctgaaa 180
gttcatatct ggagcctgat gtcttaacga ataaaggtcc catgctccac ccgaaaaaaa 240
aaaaaaaa 248

<210> 78
<211> 201
<212> DNA
<213> Homo sapien

<400> 78
actagtccag tgtggtggaa ttccattgtg ttgggcccaa cacaatggct acctttaaca 60
tcacccagac cccgccctgc ccgtgcccca cgctgctgct aacgacagta tgatgcttac 120
tctgctactc ggaaactatt tttatgtaat taatgtatgc tttcttgttt ataaatgcct 180
gatttaaaaa aaaaaaaaaa a 201

<210> 79
<211> 552
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(552)
<223> n = A,T,C or G

<400> 79
tccttttgtt aggtttttga gacaacccta gacctaaact gtgtcacaga cttctgaatg 60
tttaggcagt gctagtaatt tcctcgtaat gattctgtta ttactttcct attctttatt 120
cctctttctt ctgaagatta atgaagttga aaattgaggt ggataaatac aaaaaggtag 180
tgtgatagta taagtatcta agtgcagatg aaagtgtgtt atatatatcc attcaaaatt 240
atgcaagtta gtaattactc agggttaact aaattacttt aatatgctgt tgaacctact 300
ctgttccttg gctagaaaaa attataaaca ggactttgtt agtttgggaa gccaaattga 360
taatattcta tgttctaaaa gttgggctat acataaanta tnaagaaata tggaatttta 420
ttcccaggaa tatggggttc atttatgaat antacccggg anagaagttt tgantnaaac 480
cngttttggt taatacgtta atatgtcctn aatnaacaag gcntgactta tttccaaaaa 540
aaaaaaaaaa aa 552

<210> 80
<211> 476
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(476)
<223> n = A,T,C or G

<400> 80
acagggattt gagatgctaa ggccccagag atcgtttgat ccaaccctct tattttcaga 60
ggggaaaatg gggcctagaa gttacagagc atctagctgg tgcgctggca cccctggcct 120
cacacagact cccgagtagc tgggactaca ggcacacagt cactgaagca ggccctgttt 180
gcaattcacg ttgccacctc caacttaaac attcttcata tgtgatgtcc ttagtcacta 240
aggttaaact ttcccaccca gaaaaggcaa cttagataaa atcttagagt actttcatac 300
tcttctaagt cctcttccag cctcactttg agtcctcctt gggggttgat aggaantntc 360
tcttggcttt ctcaataaaa tctctatcca tctcatgttt aatttggtac gcntaaaaat 420
gctgaaaaaa ttaaaatgtt ctggtttcnc tttaaaaaaa aaaaaaaaaa aaaaaa 476

<210> 81
<211> 232
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(232)
<223> n = A,T,C or G

<400> 81
ttttttttttg tatgccntcn ctgtggngtt attgttgctg ccaccctgga ggagcccagt 60
ttcttctgta tctttctttt ctgggggatc ttcctggctc tgcccctcca ttcccagcct 120
ctcatcccca tcttgcactt ttgctagggt tggaggcgct ttcctggtag cccctcagag 180
actcagtcag cgggaataag tcctaggggt ggggggtgtg gcaagccggc ct 232

<210> 82
<211> 383
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(383)
<223> n = A,T,C or G

<400> 82
aggcgggagc agaagctaaa gccaaagccc aagaagagtg gcagtgccag cactggtgcc 60
agtaccagta ccaataacat gccagtgcca gtgccagcac cagtggtggc ttcagtgctg 120
gtgccagcct gaccgccact ctcacatttg ggctcttcgc tggccttggt ggagctggtg 180
ccagcaccag tggcagctct ggtgcctgtg gtttctccta caagtgagat tttagatatt 240
gttaatcctg ccagtctttc tcttcaagcc agggtgcatc ctcagaaacc tactcaacac 300
agcactctng gcagccacta tcaatcaatt gaagttgaca ctctgcatta aatctatttg 360
ccatttcaaa aaaaaaaaaa aaa 383

<210> 83
<211> 494
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(494)
<223> n = A,T,C or G

<400> 83
accgaattgg gaccgctggc ttataagcga tcatgtcctc cagtattacc tcaacgagca 60
gggagatcga gtctatacgc tgaagaaatt tgacccgatg ggacaacaga cctgctcagc 120
ccatcctgct cggttctccc cagatgacaa atactctcga caccgaatca ccatcaagaa 180
acgcttcaag gtgctcatga cccagcaacc gcgccctgtc ctctgagggt ccttaaactg 240
atgtcttttc tgccacctgt taccctcgg agactccgta accaaactct tcggactgtg 300
agccctgatg ccttttgcc agccatactc tttggcntcc agtctctcgt ggcgattgat 360
tatgcttgtg tgaggcaatc atggtggcat cacccatnaa gggaacacat ttgantttt 420
tttcncatat tttaaattac naccagaata nttcagaata aatgaattga aaaactctta 480
aaaaaaaaaa aaaa 494

<210> 84
<211> 380
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(380)
<223> n = A,T,C or G

<400> 84
gctggtagcc tatggcgtgg ccacggangg gctcctgagg cacgggacag tgacttccca 60
agtatcctgc gccgcgtctt ctaccgtccc tacctgcaga tcttcgggca gattccccag 120
gaggacatgg acgtggccct catggagcac agcaactgct cgtcggagcc cggcttctgg 180
gcacaccctc ctggggccca ggcgggcacc tgcgtctccc agtatgccaa ctggctggtg 240
gtgctgctcc tcgtcatctt cctgctcgtg gccaacatcc tgctggtcac ttgctcattg 300
ccatgttcag ttacacattc ggcaaagtac agggcaacag cnatctctac tgggaaggcc 360
agcgttnccg cctcatccgg 380

<210> 85
<211> 481
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(481)
<223> n = A,T,C or G

<400> 85
gagttagctc ctccacaacc ttgatgaggt cgtctgcagt ggcctctcgc ttcataccgc 60
tnccatcgtc atactgtagg tttgccacca cctcctgcat cttggggcgg ctaatatcca 120
ggaaactctc aatcaagtca ccgtcnatna aacctgtggc tggttctgtc ttccgctcgg 180
tgtgaaagga tctccagaag gagtgctcga tcttccccac acttttgatg actttattga 240
gtcgattctg catgtccagc aggaggttgt accagctctc tgacagtgag gtcaccagcc 300
ctatcatgcc nttgaacgtg ccgaagaaca ccgagccttg tgtggggggt gnagtctcac 360
ccagattctg cattaccaga nagccgtggc aaaaganatt gacaactcgc ccaggnngaa 420
aaagaacacc tcctggaagt gctngccgct cctcgtccnt tggtggnngc gcntnccttt 480
t 481

<210> 86
<211> 472
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(472)
<223> n = A,T,C or G

<400> 86
aacatcttcc tgtataatgc tgtgtaatat cgatccgatn ttgtctgctg agaattcatt 60
acttggaaaa gcaacttnaa gcctggacac tggtattaaa attcacaata tgcaacactt 120
taaacagtgt gtcaatctgc tcccttactt tgtcatcacc agtctgggaa taagggtatg 180
ccctattcac acctgttaaa agggcgctaa gcatttttga ttcaacatct tttttttga 240
cacaagtccg aaaaaagcaa aagtaaacag ttnttaattt gttagccaat tcactttctt 300
catgggacag agccatttga tttaaaaagc aaattgcata atattgagct ttgggagctg 360
atatntgagc ggaagantag cctttctact tcaccagaca caactccttt catattggga 420
tgttnacnaa agttatgtct cttacagatg ggatgctttt gtggcaattc tg 472

<210> 87
<211> 413
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(413)
<223> n = A,T,C or G

<400> 87
agaaaccagt atctctnaaa acaacctctc ataccttgtg gacctaattt tgtgtgcgtg 60
tgtgtgtgcg cgcatattat atagacaggc acatcttttt tacttttgta aaagcttatg 120
cctctttggt atctatatct gtgaaagttt taatgatctg ccataatgtc ttggggacct 180
ttgtcttctg tgtaaatggt actagagaaa acacctatnt tatgagtcaa tctagttngt 240
tttattcgac atgaaggaaa tttccagatn acaacactna caaactctcc cttgactagg 300

ggggacaaag aaaagcanaa ctgaacatna gaaacaattn cctggtgaga aattncataa 360
acagaaattg ggtngtatat tgaaananng catcattnaa acgtttttt ttt 413

<210> 88
<211> 448
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(448)
<223> n = A,T,C or G

<400> 88
cgcagcgggt cctctctatc tagctccagc ctctcgcctg ccccactccc cgcgtcccgc 60
gtcctagccn accatggccg ggccccctgcg cgccccgctg ctcctgctgg ccatcctggc 120
cgtggccctg gccgtgagcc ccgcggccgg ctccagtccc ggcaagccgc cgcgcctggt 180
gggaggccca tggaccccgc gtggaagaag aaggtgtgcg gcgtgcactg gactttgccg 240
tcggcnanta caacaaaccc gcaacnactt ttaccnagcn cgcgctgcag gttgtgccgc 300
cccaancaaa ttgttactng gggtaantaa ttcttggaag ttgaacctgg gccaaacnng 360
tttaccagaa ccnagccaat tngaacaatt nccccctccat aacagcccct tttaaaaagg 420
gaancantcc tgntcttttc caaatttt 448

<210> 89
<211> 463
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(463)
<223> n = A,T,C or G

<400> 89
gaattttgtg cactggccac tgtgatggaa ccattgggcc aggatgcttt gagtttatca 60
gtagtgattc tgccaaagtt ggtgttgtaa catgagtatg taaaatgtca aaaaattagc 120
agaggtctag gtctgcatat cagcagacag tttgtccgtg tattttgtag ccttgaagtt 180
ctcagtgaca agttnnttct gatgcgaagt tctnattcca gtgttttagt cctttgcatc 240
tttnatgttn agacttgcct ctntnaaatt gcttttgtnt tctgcaggta ctatctgtgg 300
tttaacaaaa tagaannact tctctgcttn gaanatttga atatcttaca tctnaaaatn 360
aattctctcc ccatannaaa acccangccc ttggganaat ttgaaaaang gntccttcnn 420
aattcnnana anttcagntn tcatacaaca naacngganc ccc 463

<210> 90
<211> 400
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(400)
<223> n = A,T,C or G

<400> 90
agggattgaa ggtctnttnt actgtcggac tgttcancca ccaactctac aagttgctgt 60
cttccactca ctgtctgtaa gcntnttaac ccagactgta tcttcataaa tagaacaaat 120
tcttcaccag tcacatcttc taggaccttt ttggattcag ttagtataag ctcttccact 180
tcctttgtta agacttcatc tggtaaagtc ttaagttttg tagaaaggaa tttaattgct 240

cgttctctaa caatgtcctc tccttgaagt atttggctga acaacccacc tnaagtccct 300
ttgtgcatcc attttaaata tacttaatag ggcattggtn cactaggtta aattctgcaa 360
gagtcatctg tctgcaaaag ttgcgttagt atatctgcca 400

<210> 91
<211> 480
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(480)
<223> n = A,T,C or G

<400> 91
gagctcggat ccaataatct ttgtctgagg gcagcacaca tatncagtgc catggnaact 60
ggtctacccc acatgggagc agcatgccgt agntatataa ggtcattccc tgagtcagac 120
atgcctcttt gactaccgtg tgccagtgct ggtgattctc acacacctcc nnccgctctt 180
tgtggaaaaa ctggcacttg nctggaacta gcaagacatc acttacaaat tcacccacga 240
gacacttgaa aggtgtaaca aagcgactct tgcattgctt tttgtccctc cggcaccagt 300
tgtcaatact aacccgctgg tttgcctcca tcacatttgt gatctgtagc tctggataca 360
tctcctgaca gtactgaaga acttcttctt ttgtttcaaa agcaactctt ggtgcctgtt 420
ngatcaggtt cccatttccc agtccgaatg ttcacatggc atatnttact tcccacaaaa 480

<210> 92
<211> 477
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(477)
<223> n = A,T,C or G

<400> 92
atacagccca natcccacca cgaagatgcg cttgttgact gagaacctga tgcggtcact 60
ggtcccgctg tagccccagc gactctccac ctgctggaag cggttgatgc tgcactcctt 120
cccacgcagg cagcagcggg gccggtcaat gaactccact cgtggcttgg ggttgacggt 180
taantgcagg aagaggctga ccacctcgcg gtccaccagg atgcccgact gtgcgggacc 240
tgcagcgaaa ctcctcgatg gtcatgagcg ggaagcgaat gangcccagg gccttgccca 300
gaaccttccg cctgttctct ggcgtcacct gcagctgctg ccgctnacac tcggcctcgg 360
accagcggac aaacggcgtt gaacagccgc acctcacgga tgcccantgt gtcgcgctcc 420
aggaacggcn ccagcgtgtc caggtcaatg tcggtgaanc ctccgcgggt aatggcg 477

<210> 93
<211> 377
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(377)
<223> n = A,T,C or G

<400> 93
gaacggctgg accttgcctc gcattgtgct gctggcagga ataccttggc aagcagctcc 60
agtccgagca gccccagacc gctgccgccc gaagctaagc ctgcctctgg ccttcccctc 120
cgcctcaatg cagaaccant agtgggagca ctgtgtttag agttaagagt gaacactgtn 180

tgattttact tgggaatttc ctctgttata tagcttttcc caatgctaat ttccaaacaa 240
caacaacaaa ataacatgtt tgcctgttna gttgtataaa agtangtgat tctgtatnta 300
aagaaaatat tactgttaca tatactgctt gcaanttctg tatttattgg tnctctggaa 360
ataaatatat tattaaa 377

<210> 94
<211> 495
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(495)
<223> n = A,T,C or G

<400> 94
ccctttgagg ggttagggtc cagttcccag tggaagaaac aggccaggag aantgcgtgc 60
cgagctgang cagatttccc acagtgaccc cagagccctg ggctatagtc tctgacccct 120
ccaaggaaag accaccttct ggggacatgg gctggagggc aggacctaga ggcaccaagg 180
gaaggcccca ttccggggct gttccccgag gaggaaggga aggggctctg tgtgcccccc 240
acgaggaana ggccctgant cctgggatca nacacccctt cacgtgtatc cccacacaaa 300
tgcaagctca ccaaggtccc ctctcagtcc cttccctaca ccctgaacgg ncactggccc 360
acacccaccc agancancca cccgccatgg ggaatgtnct caaggaatcg cngggcaacg 420
tggactctng tcccnnaagg gggcagaatc tccaatagan gganngaacc cttgctnana 480
aaaaaaaana aaaaa 495

<210> 95
<211> 472
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(472)
<223> n = A,T,C or G

<400> 95
ggttacttgg tttcattgcc accacttagt ggatgtcatt tagaaccatt ttgtctgctc 60
cctctggaag ccttgcgcag agcggacttt gtaattgttg gagaataact gctgaatttt 120
tagctgtttt gagttgattc gcaccactgc accacaactc aatatgaaaa ctatttnact 180
tatttattat cttgtgaaaa gtatacaatg aaaatttttgt tcatactgta tttatcaagt 240
atgatgaaaa gcaatagata tatattcttt tattatgttn aattatgatt gccattatta 300
atcggcaaaa tgtggagtgt atgttctttt cacagtaata tatgcctttt gtaacttcac 360
ttggttattt tattgtaaat gaattacaaa attcttaatt taagaaaatg gtangttata 420
tttanttcan taatttcttt ccttgtttac gttaattttg aaaagaatgc at 472

<210> 96
<211> 476
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(476)
<223> n = A,T,C or G

<400> 96
ctgaagcatt tcttcaaact tntctacttt tgtcattgat acctgtagta agttgacaat 60

gtggtgaaat ttcaaaatta tatgtaactt ctactagttt tactttctcc cccaagtctt 120
ttttaactca tgatttttac acacacaatc cagaacttat tatatagcct ctaagtcttt 180
attcttcaca gtagatgatg aaagagtcct ccagtgtctt gngcanaatg ttctagntat 240
agctggatac atacngtggg agttctataa actcatacct cagtgggact naaccaaaat 300
tgtgttagtc tcaattccta ccacactgag ggagcctccc aaatcactat attcttatct 360
gcaggtactc ctccagaaaa acngacaggg caggcttgca tgaaaaagtn acatctgcgt 420
tacaaagtct atcttcctca nangtctgtn aaggaacaat ttaatcttct agcttt 476

<210> 97
<211> 479
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(479)
<223> n = A,T,C or G

<400> 97
actctttcta atgctgatat gatcttgagt ataagaatgc atatgtcact agaatggata 60
aaataatgct gcaaacttaa tgttcttatg caaaatggaa cgctaatgaa acacagctta 120
caatcgcaaa tcaaaactca caagtgctca tctgttgtag atttagtgta ataagactta 180
gattgtgctc cttcggatat gattgtttct canatcttgg gcaatnttcc ttagtcaaat 240
caggctacta gaattctgtt attggatatn tgagagcatg aaatttttaa naatacactt 300
gtgattatna aattaatcac aaatttcact tatacctgct atcagcagct agaaaaacat 360
ntnntttta natcaaagta ttttgtgttt ggaantgtnn aaatgaaatc tgaatgtggg 420
ttcnatctta ttttttcccn gacnactant tnctttttta gggnctattc tganccatc 479

<210> 98
<211> 461
<212> DNA
<213> Homo sapien

<400> 98
agtgacttgt cctccaacaa aacccccttga tcaagtttgt ggcactgaca atcagaccta 60
tgctagttcc tgtcatctat tcgctactaa atgcagactg gaggggacca aaaaggggca 120
tcaactccag ctggattatt ttggagcctg caaatctatt cctacttgta cggactttga 180
agtgattcag tttcctctac ggatgagaga ctggctcaag aatatcctca tgcagcttta 240
tgaagccact ctgaacacgc tggttatcta gatgagaaca gagaaataaa gtcagaaaat 300
ttacctggag aaaagaggct ttggctgggg accatcccat tgaaccttct cttaaggact 360
ttaagaaaaa ctaccacatg ttgtgtatcc tggtgccggc cgtttatgaa ctgaccaccc 420
tttggaataa tcttgacgct cctgaacttg ctcctctgcg a 461

<210> 99
<211> 171
<212> DNA
<213> Homo sapien

<400> 99
gtggccgcgc gcaggtgttt cctcgtaccg cagggccccc tcccttcccc aggcgtccct 60
cggcgcctct gcgggcccga ggaggagcgg ctggcgggtg gggggagtgt gacccaccct 120
cggtgagaaa agccttctct agcgatctga gaggcgtgcc ttgggggtac c 171

<210> 100
<211> 269
<212> DNA
<213> Homo sapien

<400> 100
cggccgcaag tgcaactcca gctggggccg tgcggacgaa gattctgcca gcagttggtc 60
cgactgcgac gacggcggcg gcgacagtcg caggtgcagc gcgggcgcct ggggtcttgc 120
aaggctgagc tgacgccgca gaggtcgtgt cacgtcccac gaccttgacg ccgtcgggga 180
cagccggaac agagcccggt gaagcgggag gcctcgggga gcccctcggg aagggcggcc 240
cgagagatac gcaggtgcag gtggccgcc 269

<210> 101
<211> 405
<212> DNA
<213> Homo sapien

<400> 101
ttttttttttt ttttggaatc tactgcgagc acagcaggtc agcaacaagt ttattttgca 60
gctagcaagg taacagggta gggcatggtt acatgttcag gtcaacttcc tttgtcgtgg 120
ttgattggtt tgtctttatg ggggcggggt ggggtagggg aaacgaagca aataacatgg 180
agtgggtgca ccctccctgt agaacctggt tacaaagctt ggggcagttc acctggtctg 240
tgaccgtcat tttcttgaca tcaatgttat tagaagtcag gatatctttt agagagtcca 300
ctgttctgga gggagattag ggtttcttgc caaatccaac aaaatccact gaaaaagttg 360
gatgatcagt acgaataccg aggcatattc tcatatcggt ggcca 405

<210> 102
<211> 470
<212> DNA
<213> Homo sapien

<400> 102
ttttttttttt ttttttttttt ttttttttttt ttttttttttt ttttttttttt 60
ggcacttaat ccatttttat ttcaaaatgt ctacaaattt aatcccatta tacggtattt 120
tcaaaatcta aattattcaa attagccaaa tccttaccaa ataataccca aaaatcaaaa 180
atatacttct ttcagcaaac ttgttacata aattaaaaaa atatatacgg ctggtgtttt 240
caaagtacaa ttatcttaac actgcaaaca ttttaaggaa ctaaaataaa aaaaaacact 300
ccgcaaaggt taaagggaac aacaaattct tttacaacac cattataaaa atcatatctc 360
aaatcttagg ggaatatata cttcacacgg gatcttaact tttactcact ttgtttattt 420
ttttaaacca ttgtttgggc ccaacacaat ggaatccccc ctggactagt 470

<210> 103
<211> 581
<212> DNA
<213> Homo sapien

<400> 103
ttttttttttt ttttttttga ccccccctctt ataaaaaaca agttaccatt ttattttact 60
tacacatatt tattttataa ttggtattag atattcaaaa ggcagctttt aaaatcaaac 120
taaatggaaa ctgccttaga tacataattc ttaggaatta gcttaaaatc tgcctaaagt 180
gaaaatcttc tctagctctt ttgactgtaa atttttgact cttgtaaaac atccaaattc 240
atttttcttg tctttaaaat tatctaatct ttccattttt tccctattcc aagtcaattt 300
gcttctctag cctcatttcc tagctcttat ctactattag taagtggctt ttttcctaaa 360
agggaaaaca ggaagagaaa tggcacacaa aacaaacatt ttatattcat atttctacct 420
acgttaataa aatagcattt tgtgaagcca gctcaaaaga aggcttagat ccttttatgt 480
ccattttagt cactaaacga tatcaaagtg ccagaatgca aaaggtttgt gaacatttat 540
tcaaaagcta atataagata tttcacatac tcatctttct g 581

<210> 104
<211> 578
<212> DNA
<213> Homo sapien

<400> 104
ttttttttttt ttttttttttt ttttctctt ctttttttttt gaaatgagga tcgagttttt 60
cactctctag atagggcatg aagaaaactc atctttccag ctttaaaata acaatcaaat 120
ctcttatgct atatcatatt ttaagttaaa ctaatgagtc actggcttat cttctcctga 180
aggaaatctg ttcattcttc tcattcatat agttatatca agtactacct tgcatattga 240
gaggtttttc ttctctattt acacatatat ttccatgtga atttgtatca aacctttatt 300
ttcatgcaaa ctagaaaata atgtttcttt tgcataagag aagagaacaa tatagcatta 360
caaaactgct caaattgttt gttaagttat ccattataat tagttggcag gagctaatac 420
aaatcacatt tacgacagca ataataaaac tgaagtacca gttaaatatc caaaataatt 480
aaaggaacat ttttagcctg ggtataatta gctaattcac tttacaagca tttattagaa 540
tgaattcaca tgttattatt cctagcccaa cacaatgg 578

<210> 105
<211> 538
<212> DNA
<213> Homo sapien

<400> 105
ttttttttttt tttttcagta ataatcagaa caatatttat ttttatattt aaaattcata 60
gaaaagtgcc ttacatttaa taaaagtttg tttctcaaag tgatcagagg aattagatat 120
gtcttgaaca ccaatattaa tttgaggaaa atacaccaaa atacattaag taaattattt 180
aagatcatag agcttgtaag tgaaaagata aaatttgacc tcagaaactc tgagcattaa 240
aaatccacta ttagcaaata aattactatg gacttcttgc tttaattttg tgatgaatat 300
ggggtgtcac tggtaaacca acacattctg aaggatacat tacttagtga tagattctta 360
tgtactttgc taatacgtgg atatgagttg acaagtttct ctttcttcaa tcttttaagg 420
ggcgagaaat gaggaagaaa agaaaaggat tacgcatact gttctttcta tggaaggatt 480
agatatgttt cctttgccaa tattaaaaaa ataataatgt ttactactag tgaaaccc 538

<210> 106
<211> 473
<212> DNA
<213> Homo sapien

<400> 106
ttttttttttt ttttttagtc aagtttctat ttttattata attaaagtct tggtcatttc 60
atttattagc tctgcaactt acatatttaa attaaagaaa cgttttagac aactgtacaa 120
tttataaatg taaggtgcca ttattgagta atatattcct ccaagagtgg atgtgtccct 180
tctcccacca actaatgaac agcaacatta gtttaatttt attagtagat atacactgct 240
gcaaacgcta attctcttct ccatccccat gtgatattgt gtatatgtgt gagttggtag 300
aatgcatcac aatctacaat caacagcaag atgaagctag gctgggcttt cggtgaaaat 360
agactgtgtc tgtctgaatc aaatgatctg acctatcctc ggtggcaaga actcttcgaa 420
ccgcttcctc aaaggcgctg ccacatttgt ggctctttgc acttgtttca aaa 473

<210> 107
<211> 1621
<212> DNA
<213> Homo sapien

<400> 107
cgccatggca ctgcagggca tctcggtcat ggagctgtcc ggcctggccc cgggcccgtt 60
ctgtgctatg gtcctggctg acttcggggc gcgtgtggta cgcgtggacc ggcccggctc 120
ccgctacgac gtgagccgct tgggccgggg caagcgctcg ctagtgctgg acctgaagca 180
gccgcgggga gccgccgtgc tgcggcgtct gtgcaagcgg tcggatgtgc tgctggagcc 240
cttccgccgc ggtgtcatgg agaaactcca gctgggccca gagattctgc agcgggaaaa 300
tccaaggctt atttatgcca ggctgagtgg atttggccag tcaggaagct tctgccggtt 360
agctggccac gatatcaact atttggcttt gtcaggtgtt ctctcaaaaa ttggcagaag 420
tggtgagaat ccgtatgccc cgctgaatct cctggctgac tttgctggtg gtggccttat 480
gtgtgcactg ggcattataa tggctctttt tgaccgcaca cgcactgaca agggtcaggt 540

```
cattgatgca aatatggtgg aaggaacagc atatttaagt tcttttctgt ggaaaactca     600
gaaatcgagt ctgtgggaag cacctcgagg acagaacatg ttggatggtg gagcaccttt     660
ctatacgact tacaggacag cagatgggga attcatggct gttggagcaa tagaacccca     720
gttctacgag ctgctgatca aaggacttgg actaaagtct gatgaacttc ccaatcagat     780
gagcatggat gattggccag aaatgaagaa gaagtttgca gatgtatttg caaagaagac     840
gaaggcagag tggtgtcaaa tctttgacgg cacagatgcc tgtgtgactc cggttctgac     900
ttttgaggag gttgttcatc atgatcacaa caaggaacgg ggctcgttta tcaccagtga     960
ggagcaggac gtgagccccc gccctgcacc tctgctgtta aacaccccag ccatcccttc    1020
tttcaaaagg gatcctttca taggagaaca cactgaggag atacttgaag aatttggatt    1080
cagccgcgaa gagatttatc agcttaactc agataaaatc attgaaagta ataaggtaaa    1140
agctagtctc taacttccag gcccacggct caagtgaatt tgaatactgc atttacagtg    1200
tagagtaaca cataacattg tatgcatgga aacatggagg aacagtatta cagtgtccta    1260
ccactctaat caagaaaaga attacagact ctgattctac agtgatgatt gaattctaaa    1320
aatggttatc attagggctt ttgatttata aaactttggg tacttatact aaattatggt    1380
agttattctg ccttccagtt tgcttgatat atttgttgat attaagattc ttgacttata    1440
ttttgaatgg gttctagtga aaaaggaatg atatattctt gaagacatcg atatacattt    1500
atttacactc ttgattctac aatgtagaaa atgaggaaat gccacaaatt gtatggtgat    1560
aaaagtcacg tgaaacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa    1620
a                                                                    1621
```

<210> 108
<211> 382
<212> PRT
<213> Homo sapien

<400> 108

```
Met Ala Leu Gln Gly Ile Ser Val Met Glu Leu Ser Gly Leu Ala Pro
 1               5                  10                  15
Gly Pro Phe Cys Ala Met Val Leu Ala Asp Phe Gly Ala Arg Val Val
            20                  25                  30
Arg Val Asp Arg Pro Gly Ser Arg Tyr Asp Val Ser Arg Leu Gly Arg
        35                  40                  45
Gly Lys Arg Ser Leu Val Leu Asp Leu Lys Gln Pro Arg Gly Ala Ala
    50                  55                  60
Val Leu Arg Arg Leu Cys Lys Arg Ser Asp Val Leu Leu Glu Pro Phe
65                  70                  75                  80
Arg Arg Gly Val Met Glu Lys Leu Gln Leu Gly Pro Glu Ile Leu Gln
                85                  90                  95
Arg Glu Asn Pro Arg Leu Ile Tyr Ala Arg Leu Ser Gly Phe Gly Gln
            100                 105                 110
Ser Gly Ser Phe Cys Arg Leu Ala Gly His Asp Ile Asn Tyr Leu Ala
        115                 120                 125
Leu Ser Gly Val Leu Ser Lys Ile Gly Arg Ser Gly Glu Asn Pro Tyr
    130                 135                 140
Ala Pro Leu Asn Leu Leu Ala Asp Phe Ala Gly Gly Gly Leu Met Cys
145                 150                 155                 160
Ala Leu Gly Ile Ile Met Ala Leu Phe Asp Arg Thr Arg Thr Asp Lys
                165                 170                 175
Gly Gln Val Ile Asp Ala Asn Met Val Glu Gly Thr Ala Tyr Leu Ser
            180                 185                 190
Ser Phe Leu Trp Lys Thr Gln Lys Ser Ser Leu Trp Glu Ala Pro Arg
        195                 200                 205
Gly Gln Asn Met Leu Asp Gly Gly Ala Pro Phe Tyr Thr Thr Tyr Arg
    210                 215                 220
Thr Ala Asp Gly Glu Phe Met Ala Val Gly Ala Ile Glu Pro Gln Phe
225                 230                 235                 240
Tyr Glu Leu Leu Ile Lys Gly Leu Gly Leu Lys Ser Asp Glu Leu Pro
                245                 250                 255
```

```
Asn Gln Met Ser Met Asp Asp Trp Pro Glu Met Lys Lys Lys Phe Ala
            260                 265                 270
Asp Val Phe Ala Lys Lys Thr Lys Ala Glu Trp Cys Gln Ile Phe Asp
        275                 280                 285
Gly Thr Asp Ala Cys Val Thr Pro Val Leu Thr Phe Glu Glu Val Val
    290                 295                 300
His His Asp His Asn Lys Glu Arg Gly Ser Phe Ile Thr Ser Glu Glu
305                 310                 315                 320
Gln Asp Val Ser Pro Arg Pro Ala Pro Leu Leu Leu Asn Thr Pro Ala
                325                 330                 335
Ile Pro Ser Phe Lys Arg Asp Pro Phe Ile Gly Glu His Thr Glu Glu
            340                 345                 350
Ile Leu Glu Glu Phe Gly Phe Ser Arg Glu Glu Ile Tyr Gln Leu Asn
        355                 360                 365
Ser Asp Lys Ile Ile Glu Ser Asn Lys Val Lys Ala Ser Leu
    370                 375                 380
```

<210> 109
<211> 1524
<212> DNA
<213> Homo sapien

<400> 109

```
ggcacgaggc tgcgccaggg cctgagcgga ggcgggggca gcctcgccag cgggggcccc      60
gggcctggcc atgcctcact gagccagcgc ctgcgcctct acctcgccga cagctggaac     120
cagtgcgacc tagtggctct cacctgcttc ctcctgggcg tgggctgccg gctgaccccg     180
ggtttgtacc acctgggccg cactgtcctc tgcatcgact tcatggtttt cacggtgcgg     240
ctgcttcaca tcttcacggt caacaaacag ctggggccca agatcgtcat cgtgagcaag     300
atgatgaagg acgtgttctt cttcctcttc ttcctcggcg tgtggctggt agcctatggc     360
gtggccacgg aggggctcct gaggccacgg gacagtgact tcccaagtat cctgcgccgc     420
gtcttctacc gtccctacct gcagatcttc gggcagattc cccaggagga catggacgtg     480
gccctcatgg agcacagcaa ctgctcgtcg gagcccggct tctgggcaca ccctcctggg     540
gcccaggcgg gcacctgcgt ctcccagtat gccaactggc tggtggtgct gctcctcgtc     600
atcttcctgc tcgtggccaa catcctgctg gtcaacttgc tcattgccat gttcagttac     660
acattcggca aagtacaggg caacagcgat ctctactgga aggcgcagcg ttaccgcctc     720
atccgggaat tccactctcg gcccgcgctg gccccgccct ttatcgtcat ctcccacttg     780
cgcctcctgc tcaggcaatt gtgcaggcga ccccggagcc cccagccgtc ctccccggcc     840
ctcgagcatt tccgggttta cctttctaag gaagccgagc ggaagctgct aacgtgggaa     900
tcggtgcata aggagaactt tctgctggca cgcgctaggg acaagcggga gagcgactcc     960
gagcgtctga agcgcacgtc ccagaaggtg gacttggcac tgaaacagct gggacacatc    1020
cgcgagtacg aacagcgcct gaaagtgctg gagcgggagg tccagcagtg tagccgcgtc    1080
ctggggtggg tggccgaggc cctgagccgc tctgccttgc tgcccccagg tgggccgcca    1140
ccccctgacc tgcctgggtc caaagactga gccctgctgg cggacttcaa ggagaagccc    1200
ccacagggga ttttgctcct agagtaaggc tcatctgggc ctcggccccc gcacctggtg    1260
gccttgtcct tgaggtgagc cccatgtcca tctgggccac tgtcaggacc acctttggga    1320
gtgtcatcct tacaaaccac agcatgcccg gctcctccca gaaccagtcc cagcctggga    1380
ggatcaaggc ctggatcccg ggccgttatc catctggagg ctgcagggtc cttggggtaa    1440
cagggaccac agacccctca ccactcacag attcctcaca ctggggaaat aaagccattt    1500
cagaggaaaa aaaaaaaaaa aaaa                                           1524
```

<210> 110
<211> 3410
<212> DNA
<213> Homo sapien

<400> 110

```
gggaaccagc ctgcacgcgc tggctccggg tgacagccgc gcgcctcggc caggatctga      60
gtgatgagac gtgtccccac tgaggtgccc cacagcagca ggtgttgagc atgggctgag     120
```

```
aagctggacc ggcaccaaag ggctggcaga aatgggcgcc tggctgattc ctaggcagtt     180
ggcggcagca aggaggagag gccgcagctt ctggagcaga gccgagacga agcagttctg     240
gagtgcctga acggcccccc gagccctacc cgcctggccc actatggtcc agaggctgtg     300
ggtgagccgc ctgctgcggc accggaaagc ccagctcttg ctggtcaacc tgctaacctt     360
tggcctggag gtgtgtttgg ccgcaggcat cacctatgtg ccgcctctgc tgctggaagt     420
gggggtagag gagaagttca tgaccatggt gctgggcatt ggtccagtgc tgggcctggt     480
ctgtgtcccg ctcctaggct cagccagtga ccactggcgt ggacgctatg gccgccgccg     540
gcccttcatc tgggcactgt ccttgggcat cctgctgagc ctctttctca tcccaagggc     600
cggctggcta gcagggctgc tgtgcccgga tcccaggccc ctggagctgg cactgctcat     660
cctgggcgtg gggctgctgg acttctgtgg ccaggtgtgc ttcactccac tggaggccct     720
gctctctgac ctcttccggg acccggacca ctgtcgccag gcctactctg tctatgcctt     780
catgatcagt cttgggggct gcctgggcta cctcctgcct gccattgact gggacaccag     840
tgccctggcc ccctacctgg gcacccagga ggagtgcctc tttggcctgc tcaccctcat     900
cttcctcacc tgcgtagcag ccacactgct ggtggctgag gaggcagcgc tgggccccac     960
cgagccagca gaagggctgt cggccccctc cttgtcgccc cactgctgtc catgccgggc    1020
ccgcttggct ttccggaacc tgggcgccct gcttccccgg ctgcaccagc tgtgctgccg    1080
catgccccgc accctgcgcc ggctcttcgt ggctgagctg tgcagctgga tggcactcat    1140
gaccttcacg ctgttttaca cggatttcgt gggcgagggg ctgtaccagg gcgtgcccag    1200
agctgagccg ggcaccgagg cccggagaca ctatgatgaa ggcgttcgga tgggcagcct    1260
ggggctgttc ctgcagtgcg ccatctccct ggtcttctct ctggtcatgg accggctggt    1320
gcagcgattc ggcactcgag cagtctattt ggccagtgtg gcagctttcc ctgtggctgc    1380
cggtgccaca tgcctgtccc acagtgtggc cgtggtgaca gcttcagccg ccctcaccgg    1440
gttcaccttc tcagccctgc agatcctgcc ctacacactg gcctccctct accaccggga    1500
gaagcaggtg ttcctgccca aataccgagg ggacactgga ggtgctagca gtgaggacag    1560
cctgatgacc agcttcctgc caggccctaa gcctggagct cccttcccta atggacacgt    1620
gggtgctgga ggcagtggcc tgctcccacc tccacccgcg ctctgcgggg cctctgcctg    1680
tgatgtctcc gtacgtgtgg tggtgggtga gcccaccgag gccagggtgg ttccgggccg    1740
gggcatctgc ctggacctcg ccatcctgga tagtgccttc ctgctgtccc aggtggcccc    1800
atccctgttt atgggctcca ttgtccagct cagccagtct gtcactgcct atatggtgtc    1860
tgccgcaggc ctgggtctgg tcgccattta ctttgctaca caggtagtat ttgacaagag    1920
cgacttggcc aaatactcag cgtagaaaac ttccagcaca ttggggtgga gggcctgcct    1980
cactgggtcc cagctccccg ctcctgttag ccccatgggg ctgccgggct ggccgccagt    2040
ttctgttgct gccaaagtaa tgtggctctc tgctgccacc ctgtgctgct gaggtgcgta    2100
gctgcacagc tgggggctgg ggcgtccctc tcctctctcc ccagtctcta gggctgcctg    2160
actggaggcc ttccaagggg gtttcagtct ggacttatac agggaggcca gaagggctcc    2220
atgcactgga atgcggggac tctgcaggtg gattacccag gctcagggtt aacagctagc    2280
ctcctagttg agacacacct agagaagggt ttttgggagc tgaataaact cagtcacctg    2340
gtttcccatc tctaagcccc ttaacctgca gcttcgttta atgtagctct tgcatgggag    2400
tttctaggat gaaacactcc tccatgggat ttgaacatat gacttatttg taggggaaga    2460
gtcctgaggg gcaacacaca agaaccaggt cccctcagcc cacagcactg tctttttgct    2520
gatccacccc cctcttacct tttatcagga tgtggcctgt tggtccttct gttgccatca    2580
cagagacaca ggcatttaaa tatttaactt atttatttaa caaagtagaa gggaatccat    2640
tgctagcttt tctgtgttgg tgtctaatat ttgggtaggg tgggggatcc ccaacaatca    2700
ggtcccctga gatagctggt cattgggctg atcattgcca gaatcttctt ctcctggggt    2760
ctggcccccc aaaatgccta acccaggacc ttggaaattc tactcatccc aaatgataat    2820
tccaaatgct gttacccaag gttagggtgt tgaaggaagg tagagggtgg ggcttcaggt    2880
ctcaacggct tccctaacca cccctcttct cttggcccag cctggttccc cccacttcca    2940
ctcccctcta ctctctctag gactgggctg atgaaggcac tgcccaaaat ttcccctacc    3000
cccaactttc ccctaccccc aactttcccc accagctcca caaccctgtt tggagctact    3060
gcaggaccag aagcacaaag tgcggtttcc caagcctttg tccatctcag cccccagagt    3120
atatctgtgc ttggggaatc tcacacagaa actcaggagc acccccctgcc tgagctaagg    3180
gaggtcttat ctctcagggg gggtttaagt gccgtttgca ataatgtcgt cttatttatt    3240
tagcggggtg aatattttat actgtaagtg agcaatcaga gtataatgtt tatggtgaca    3300
aaattaaagg ctttcttata tgtttaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa    3360
aaaaaaaara aaaaaaaaaa aaaaaaaaaa aaaaaaataa aaaaaaaaaa               3410
```

<210> 111
<211> 1289

<212> DNA
<213> Homo sapien

<400> 111
agccaggcgt ccctctgcct gcccactcag tggcaacacc cgggagctgt tttgtccttt 60
gtggagcctc agcagttccc tctttcagaa ctcactgcca agagccctga acaggagcca 120
ccatgcagtg cttcagcttc attaagacca tgatgatcct cttcaatttg ctcatctttc 180
tgtgtggtgc agccctgttg gcagtgggca tctgggtgtc aatcgatggg gcatcctttc 240
tgaagatctt cgggccactg tcgtccagtg ccatgcagtt tgtcaacgtg ggctacttcc 300
tcatcgcagc cggcgttgtg gtctttgctc ttggtttcct gggctgctat ggtgctaaga 360
ctgagagcaa gtgtgccctc gtgacgttct tcttcatcct cctcctcatc ttcattgctg 420
aggttgcagc tgctgtggtc gccttggtgt acaccacaat ggctgagcac ttcctgacgt 480
tgctggtagt gcctgccatc aagaaagatt atggttccca ggaagacttc actcaagtgt 540
ggaacaccac catgaaaggg ctcaagtgct gtggcttcac caactatacg gattttgagg 600
actcacccta cttcaaagag aacagtgcct ttcccccatt ctgttgcaat gacaacgtca 660
ccaacacagc caatgaaacc tgcaccaagc aaaaggctca cgaccaaaaa gtagagggtt 720
gcttcaatca gcttttgtat gacatccgaa ctaatgcagt caccgtgggt ggtgtggcag 780
ctggaattgg gggcctcgag ctggctgcca tgattgtgtc catgtatctg tactgcaatc 840
tacaataagt ccacttctgc ctctgccact actgctgcca catgggaact gtgaagaggc 900
accctggcaa gcagcagtga ttgggggagg ggacaggatc taacaatgtc acttgggcca 960
gaatggacct gccctttctg ctccagactt ggggctagat agggaccact ccttttagcg 1020
atgcctgact ttccttccat tggtgggtgg atgggtgggg ggcattccag agcctctaag 1080
gtagccagtt ctgttgccca ttcccccagt ctattaaacc cttgatatgc cccctaggcc 1140
tagtggtgat cccagtgctc tactggggga tgagagaaag gcattttata gcctgggcat 1200
aagtgaaatc agcagagcct ctgggtggat gtgtagaagg cacttcaaaa tgcataaacc 1260
tgttacaatg ttaaaaaaaa aaaaaaaaa 1289

<210> 112
<211> 315
<212> PRT
<213> Homo sapien

<400> 112
Met Val Phe Thr Val Arg Leu Leu His Ile Phe Thr Val Asn Lys Gln
1 5 10 15
Leu Gly Pro Lys Ile Val Ile Val Ser Lys Met Met Lys Asp Val Phe
20 25 30
Phe Phe Leu Phe Phe Leu Gly Val Trp Leu Val Ala Tyr Gly Val Ala
35 40 45
Thr Glu Gly Leu Leu Arg Pro Arg Asp Ser Asp Phe Pro Ser Ile Leu
50 55 60
Arg Arg Val Phe Tyr Arg Pro Tyr Leu Gln Ile Phe Gly Gln Ile Pro
65 70 75 80
Gln Glu Asp Met Asp Val Ala Leu Met Glu His Ser Asn Cys Ser Ser
85 90 95
Glu Pro Gly Phe Trp Ala His Pro Pro Gly Ala Gln Ala Gly Thr Cys
100 105 110
Val Ser Gln Tyr Ala Asn Trp Leu Val Val Leu Leu Leu Val Ile Phe
115 120 125
Leu Leu Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Met Phe
130 135 140
Ser Tyr Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu Tyr Trp Lys
145 150 155 160
Ala Gln Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg Pro Ala Leu
165 170 175
Ala Pro Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu Leu Arg Gln
180 185 190
Leu Cys Arg Arg Pro Arg Ser Pro Gln Pro Ser Ser Pro Ala Leu Glu

```
        195                 200                 205
His Phe Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg Lys Leu Leu Thr
    210                 215                 220
Trp Glu Ser Val His Lys Glu Asn Phe Leu Leu Ala Arg Ala Arg Asp
225                 230                 235                 240
Lys Arg Glu Ser Asp Ser Glu Arg Leu Lys Arg Thr Ser Gln Lys Val
                245                 250                 255
Asp Leu Ala Leu Lys Gln Leu Gly His Ile Arg Glu Tyr Glu Gln Arg
            260                 265                 270
Leu Lys Val Leu Glu Arg Glu Val Gln Gln Cys Ser Arg Val Leu Gly
        275                 280                 285
Trp Val Ala Glu Ala Leu Ser Arg Ser Ala Leu Leu Pro Pro Gly Gly
    290                 295                 300
Pro Pro Pro Pro Asp Leu Pro Gly Ser Lys Asp
305                 310                 315
```

<210> 113
<211> 553
<212> PRT
<213> Homo sapien

<400> 113

```
Met Val Gln Arg Leu Trp Val Ser Arg Leu Leu Arg His Arg Lys Ala
 1               5                  10                  15
Gln Leu Leu Leu Val Asn Leu Leu Thr Phe Gly Leu Glu Val Cys Leu
            20                  25                  30
Ala Ala Gly Ile Thr Tyr Val Pro Pro Leu Leu Leu Glu Val Gly Val
        35                  40                  45
Glu Glu Lys Phe Met Thr Met Val Leu Gly Ile Gly Pro Val Leu Gly
    50                  55                  60
Leu Val Cys Val Pro Leu Leu Gly Ser Ala Ser Asp His Trp Arg Gly
65                  70                  75                  80
Arg Tyr Gly Arg Arg Arg Pro Phe Ile Trp Ala Leu Ser Leu Gly Ile
                85                  90                  95
Leu Leu Ser Leu Phe Leu Ile Pro Arg Ala Gly Trp Leu Ala Gly Leu
            100                 105                 110
Leu Cys Pro Asp Pro Arg Pro Leu Glu Leu Ala Leu Leu Ile Leu Gly
        115                 120                 125
Val Gly Leu Leu Asp Phe Cys Gly Gln Val Cys Phe Thr Pro Leu Glu
    130                 135                 140
Ala Leu Leu Ser Asp Leu Phe Arg Asp Pro Asp His Cys Arg Gln Ala
145                 150                 155                 160
Tyr Ser Val Tyr Ala Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr
                165                 170                 175
Leu Leu Pro Ala Ile Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu
            180                 185                 190
Gly Thr Gln Glu Glu Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu
        195                 200                 205
Thr Cys Val Ala Ala Thr Leu Leu Val Ala Glu Glu Ala Ala Leu Gly
    210                 215                 220
Pro Thr Glu Pro Ala Glu Gly Leu Ser Ala Pro Ser Leu Ser Pro His
225                 230                 235                 240
Cys Cys Pro Cys Arg Ala Arg Leu Ala Phe Arg Asn Leu Gly Ala Leu
                245                 250                 255
Leu Pro Arg Leu His Gln Leu Cys Cys Arg Met Pro Arg Thr Leu Arg
            260                 265                 270
Arg Leu Phe Val Ala Glu Leu Cys Ser Trp Met Ala Leu Met Thr Phe
        275                 280                 285
```

```
Thr Leu Phe Tyr Thr Asp Phe Val Gly Glu Gly Leu Tyr Gln Gly Val
    290                 295                 300
Pro Arg Ala Glu Pro Gly Thr Glu Ala Arg Arg His Tyr Asp Glu Gly
305                 310                 315                 320
Val Arg Met Gly Ser Leu Gly Leu Phe Leu Gln Cys Ala Ile Ser Leu
                325                 330                 335
Val Phe Ser Leu Val Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg
            340                 345                 350
Ala Val Tyr Leu Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala
        355                 360                 365
Thr Cys Leu Ser His Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu
    370                 375                 380
Thr Gly Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala
385                 390                 395                 400
Ser Leu Tyr His Arg Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly
                405                 410                 415
Asp Thr Gly Gly Ala Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu
            420                 425                 430
Pro Gly Pro Lys Pro Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala
        435                 440                 445
Gly Gly Ser Gly Leu Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser
    450                 455                 460
Ala Cys Asp Val Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala
465                 470                 475                 480
Arg Val Val Pro Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp
                485                 490                 495
Ser Ala Phe Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met Gly Ser
            500                 505                 510
Ile Val Gln Leu Ser Gln Ser Val Thr Ala Tyr Met Val Ser Ala Ala
        515                 520                 525
Gly Leu Gly Leu Val Ala Ile Tyr Phe Ala Thr Gln Val Val Phe Asp
    530                 535                 540
Lys Ser Asp Leu Ala Lys Tyr Ser Ala
545                 550
```

<210> 114
<211> 241
<212> PRT
<213> Homo sapien

<400> 114

```
Met Gln Cys Phe Ser Phe Ile Lys Thr Met Met Ile Leu Phe Asn Leu
 1               5                  10                  15
Leu Ile Phe Leu Cys Gly Ala Ala Leu Leu Ala Val Gly Ile Trp Val
            20                  25                  30
Ser Ile Asp Gly Ala Ser Phe Leu Lys Ile Phe Gly Pro Leu Ser Ser
        35                  40                  45
Ser Ala Met Gln Phe Val Asn Val Gly Tyr Phe Leu Ile Ala Ala Gly
    50                  55                  60
Val Val Val Phe Ala Leu Gly Phe Leu Gly Cys Tyr Gly Ala Lys Thr
65                  70                  75                  80
Glu Ser Lys Cys Ala Leu Val Thr Phe Phe Phe Ile Leu Leu Leu Ile
                85                  90                  95
Phe Ile Ala Glu Val Ala Ala Ala Val Val Ala Leu Val Tyr Thr Thr
            100                 105                 110
Met Ala Glu His Phe Leu Thr Leu Leu Val Val Pro Ala Ile Lys Lys
        115                 120                 125
Asp Tyr Gly Ser Gln Glu Asp Phe Thr Gln Val Trp Asn Thr Thr Met
```

```
    130                 135                 140
Lys Gly Leu Lys Cys Cys Gly Phe Thr Asn Tyr Thr Asp Phe Glu Asp
145                 150                 155                 160
Ser Pro Tyr Phe Lys Glu Asn Ser Ala Phe Pro Pro Phe Cys Cys Asn
                165                 170                 175
Asp Asn Val Thr Asn Thr Ala Asn Glu Thr Cys Thr Lys Gln Lys Ala
            180                 185                 190
His Asp Gln Lys Val Glu Gly Cys Phe Asn Gln Leu Leu Tyr Asp Ile
        195                 200                 205
Arg Thr Asn Ala Val Thr Val Gly Gly Val Ala Ala Gly Ile Gly Gly
    210                 215                 220
Leu Glu Leu Ala Ala Met Ile Val Ser Met Tyr Leu Tyr Cys Asn Leu
225                 230                 235                 240
Gln
```

<210> 115
<211> 366
<212> DNA
<213> Homo sapien

<400> 115

```
gctctttctc tcccctcctc tgaatttaat tctttcaact tgcaatttgc aaggattaca      60
catttcactg tgatgtatat tgtgttgcaa aaaaaaaaaa gtgtctttgt ttaaaattac     120
ttggtttgtg aatccatctt gctttttccc cattggaact agtcattaac ccatctctga     180
actggtagaa aaacatctga agagctagtc tatcagcatc tgacaggtga attggatggt     240
tctcagaacc atttcaccca gacagcctgt ttctatcctg tttaataaat tagtttgggt     300
tctctacatg cataacaaac cctgctccaa tctgtcacat aaaagtctgt gacttgaagt     360
ttagtc                                                               366
```

<210> 116
<211> 282
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(282)
<223> n = A,T,C or G

<400> 116

```
acaaagatga accatttcct atattatagc aaaattaaaa tctacccgta ttctaatatt      60
gagaaatgag atnaaacaca atnttataaa gtctacttag agaagatcaa gtgacctcaa     120
agactttact attttcatat tttaagacac atgatttatc ctattttagt aacctggttc     180
atacgttaaa caaaggataa tgtgaacagc agagaggatt tgttggcaga aaatctatgt     240
tcaatctnga actatctana tcacagacat ttctattcct tt                       282
```

<210> 117
<211> 305
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(305)
<223> n = A,T,C or G

<400> 117

acacatgtcg cttcactgcc ttcttagatg cttctggtca acatanagga acagggacca 60
tatttatcct ccctcctgaa acaattgcaa aataanacaa aatatatgaa acaattgcaa 120
aataaggcaa aatatatgaa acaacaggtc tcgagatatt ggaaatcagt caatgaagga 180
tactgatccc tgatcactgt cctaatgcag gatgtgggaa acagatgagg tcacctctgt 240
gactgcccca gcttactgcc tgtagagagt ttctangctg cagttcagac agggagaaat 300
tgggt 305

<210> 118
<211> 71
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(71)
<223> n = A,T,C or G

<400> 118
accaaggtgt ntgaatctct gacgtgggga tctctgattc ccgcacaatc tgagtggaaa 60
aantcctggg t 71

<210> 119
<211> 212
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(212)
<223> n = A,T,C or G

<400> 119
actccggttg gtgtcagcag cacgtggcat tgaacatngc aatgtggagc ccaaaccaca 60
gaaaatgggg tgaaattggc caactttcta tnaacttatg ttggcaantt tgccaccaac 120
agtaagctgg cccttctaat aaaagaaaat tgaaaggttt ctcactaanc ggaattaant 180
aatggantca aganactccc aggcctcagc gt 212

<210> 120
<211> 90
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(90)
<223> n = A,T,C or G

<400> 120
actcgttgca natcaggggc cccccagagt caccgttgca ggagtccttc tggtcttgcc 60
ctccgccggc gcagaacatg ctggggtggt 90

<210> 121
<211> 218
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature

<222> (1)...(218)
<223> n = A,T,C or G

<400> 121
tgtancgtga anacgacaga nagggttgtc aaaaatggag aanccttgaa gtcattttga 60
gaataagatt tgctaaaaga tttggggcta aaacatggtt attgggagac atttctgaag 120
atatncangt aaattangga atgaattcat ggttcttttg ggaattcctt tacgatngcc 180
agcatanact tcatgtgggg atancagcta cccttgta 218

<210> 122
<211> 171
<212> DNA
<213> Homo sapien

<400> 122
taggggtgta tgcaactgta aggacaaaaa ttgagactca actggcttaa ccaataaagg 60
catttgttag ctcatggaac aggaagtcgg atggtggggc atcttcagtg ctgcatgagt 120
caccaccccg gcggggtcat ctgtgccaca ggtccctgtt gacagtgcgg t 171

<210> 123
<211> 76
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(76)
<223> n = A,T,C or G

<400> 123
tgtagcgtga agacnacaga atggtgtgtg ctgtgctatc caggaacaca tttattatca 60
ttatcaanta ttgtgt 76

<210> 124
<211> 131
<212> DNA
<213> Homo sapien

<400> 124
acctttcccc aaggccaatg tcctgtgtgc taactggccg gctgcaggac agctgcaatt 60
caatgtgctg ggtcatatgg aggggaggag actctaaaat agccaatttt attctcttgg 120
ttaagatttg t 131

<210> 125
<211> 432
<212> DNA
<213> Homo sapien

<400> 125
actttatcta ctggctatga aatagatggt ggaaaattgc gttaccaact ataccactgg 60
cttgaaaaag aggtgatagc tcttcagagg acttgtgact tttgctcaga tgctgaagaa 120
ctacagtctg catttggcag aaatgaagat gaatttggat taaatgagga tgctgaagat 180
ttgcctcacc aaacaaaagt gaaacaactg agagaaaatt ttcaggaaaa aagacagtgg 240
ctcttgaagt atcagtcact tttgagaatg tttcttagtt actgcatact tcatggatcc 300
catggtgggg gtcttgcatc tgtaagaatg gaattgattt tgcttttgca agaatctcag 360
caggaaacat cagaaccact attttctagc cctctgtcag agcaaacctc agtgcctctc 420
ctctttgctt gt 432

<210> 126
<211> 112
<212> DNA
<213> Homo sapien

<400> 126
acacaacttg aatagtaaaa tagaaactga gctgaaattt ctaattcact ttctaaccat 60
agtaagaatg atatttcccc ccagggatca ccaaatattt ataaaaattt gt 112

<210> 127
<211> 54
<212> DNA
<213> Homo sapien

<400> 127
accacgaaac cacaaacaag atggaagcat caatccactt gccaagcaca gcag 54

<210> 128
<211> 323
<212> DNA
<213> Homo sapien

<400> 128
acctcattag taattgtttt gttgtttcat ttttttctaa tgtctcccct ctaccagctc 60
acctgagata acagaatgaa aatggaagga cagccagatt tctcctttgc tctctgctca 120
ttctctctga agtctaggtt acccattttg gggacccatt ataggcaata aacacagttc 180
ccaaagcatt tggacagttt cttgttgtgt tttagaatgg ttttcctttt tcttagcctt 240
ttcctgcaaa aggctcactc agtcccttgc ttgctcagtg gactgggctc cccagggcct 300
aggctgcctt cttttccatg tcc 323

<210> 129
<211> 192
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(192)
<223> n = A,T,C or G

<400> 129
acatacatgt gtgtatattt ttaaatatca cttttgtatc actctgactt tttagcatac 60
tgaaaacaca ctaacataat ttntgtgaac catgatcaga tacaacccaa atcattcatc 120
tagcacattc atctgtgata naaagatagg tgagtttcat ttccttcacg ttggccaatg 180
gataaacaaa gt 192

<210> 130
<211> 362
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(362)
<223> n = A,T,C or G

<400> 130
ccctttttta tggaatgagt agactgtatg tttgaanatt tanccacaac ctctttgaca 60

tataatgacg caacaaaaag gtgctgttta gtcctatggt tcagtttatg cccctgacaa 120
gtttccattg tgttttgccg atcttctggc taatcgtggt atcctccatg ttattagtaa 180
ttctgtattc cattttgtta acgcctggta gatgtaacct gctangaggc taactttata 240
cttatttaaa agctcttatt ttgtggtcat taaaatggca atttatgtgc agcactttat 300
tgcagcagga agcacgtgtg ggttggttgt aaagctcttt gctaatctta aaaagtaatg 360
gg 362

<210> 131
<211> 332
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(332)
<223> n = A,T,C or G

<400> 131
cttttgaaa gatcgtgtcc actcctgtgg acatcttgtt ttaatggagt ttcccatgca 60
gtangactgg tatggttgca gctgtccaga taaaaacatt tgaagagctc caaaatgaga 120
gttctcccag gttcgccctg ctgctccaag tctcagcagc agcctctttt aggaggcatc 180
ttctgaacta gattaaggca gcttgtaaat ctgatgtgat ttggtttatt atccaactaa 240
cttccatctg ttatcactgg agaaagccca gactccccan gacnggtacg gattgtgggc 300
atanaaggat tgggtgaagc tggcgttgtg gt 332

<210> 132
<211> 322
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(322)
<223> n = A,T,C or G

<400> 132
acttttgcca ttttgtatat ataaacaatc ttgggacatt ctcctgaaaa ctaggtgtcc 60
agtggctaag agaactcgat ttcaagcaat tctgaaagga aaaccagcat gacacagaat 120
ctcaaattcc caaacagggg ctctgtggga aaaatgaggg aggacctttg tatctcgggt 180
tttagcaagt taaaatgaan atgacaggaa aggcttattt atcaacaaag agaagagttg 240
ggatgcttct aaaaaaaact ttggtagaga aaataggaat gctnaatcct agggaagcct 300
gtaacaatct acaattggtc ca 322

<210> 133
<211> 278
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(278)
<223> n = A,T,C or G

<400> 133
acaagccttc acaagtttaa ctaaattggg attaatcttt ctgtanttat ctgcataatt 60
cttgtttttc tttccatctg gctcctgggt tgacaatttg tggaaacaac tctattgcta 120
ctatttaaaa aaaatcacaa atctttccct ttaagctatg ttnaattcaa actattcctg 180
ctattcctgt tttgtcaaag aaattatatt tttcaaaata tgtntatttg tttgatgggt 240

cccacgaaac actaataaaa accacagaga ccagcctg 278

<210> 134
<211> 121
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(121)
<223> n = A,T,C or G

<400> 134
gtttanaaaa cttgtttagc tccatagagg aaagaatgtt aaactttgta ttttaaaaca 60
tgattctctg aggttaaact tggttttcaa atgttatttt tacttgtatt ttgcttttgg 120
t 121

<210> 135
<211> 350
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(350)
<223> n = A,T,C or G

<400> 135
acttanaacc atgcctagca catcagaatc cctcaaagaa catcagtata atcctatacc 60
atancaagtg gtgactggtt aagcgtgcga caaaggtcag ctggcacatt acttgtgtgc 120
aaacttgata cttttgttct aagtaggaac tagtatacag tncctaggan tggtactcca 180
gggtgccccc caactcctgc agccgctcct ctgtgccagn ccctgnaagg aactttcgct 240
ccacctcaat caagccctgg gccatgctac ctgcaattgg ctgaacaaac gtttgctgag 300
ttcccaagga tgcaaagcct ggtgctcaac tcctggggcg tcaactcagt 350

<210> 136
<211> 399
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(399)
<223> n = A,T,C or G

<400> 136
tgtaccgtga agacgacaga agttgcatgg cagggacagg gcagggccga ggccagggtt 60
gctgtgattg tatccgaata ntcctcgtga gaaaagataa tgagatgacg tgagcagcct 120
gcagacttgt gtctgccttc aanaagccag acaggaaggc cctgcctgcc ttggctctga 180
cctggcggcc agccagccag ccacaggtgg gcttcttcct tttgtggtga caacnccaag 240
aaaactgcag aggcccaggg tcaggtgtna gtgggtangt gaccataaaa caccaggtgc 300
tcccaggaac ccgggcaaag gccatcccca cctacagcca gcatgcccac tggcgtgatg 360
ggtgcagang gatgaagcag ccagntgttc tgctgtggt 399

<210> 137
<211> 165
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(165)
<223> n = A,T,C or G

<400> 137
actggtgtgg tnggggtga tgctggtggt anaagttgan gtgacttcan gatggtgtgt 60
ggaggaagtg tgtgaacgta gggatgtaga ngttttggcc gtgctaaatg agcttcggga 120
ttggctggtc ccactggtgg tcactgtcat tggtggggtt cctgt 165

<210> 138
<211> 338
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(338)
<223> n = A,T,C or G

<400> 138
actcactgga atgccacatt cacaacagaa tcagaggtct gtgaaaacat taatggctcc 60
ttaacttctc cagtaagaat cagggacttg aaatggaaac gttaacagcc acatgcccaa 120
tgctgggcag tctcccatgc cttccacagt gaaagggctt gagaaaaatc acatccaatg 180
tcatgtgttt ccagccacac caaaaggtgc ttggggtgga gggctggggg catananggt 240
cangcctcag gaagcctcaa gttccattca gctttgccac tgtacattcc ccatntttaa 300
aaaaactgat gccttttttt ttttttttg taaaattc 338

<210> 139
<211> 382
<212> DNA
<213> Homo sapien

<400> 139
gggaatcttg gtttttggca tctggtttgc ctatagccga ggccactttg acagaacaaa 60
gaaagggact tcgagtaaga aggtgattta cagccagcct agtgcccgaa gtgaaggaga 120
attcaaacag acctcgtcat tcctggtgtg agcctggtcg gctcaccgcc tatcatctgc 180
atttgcctta ctcaggtgct accggactct ggcccctgat gtctgtagtt tcacaggatg 240
ccttatttgt cttctacacc ccacagggcc ccctacttct tcggatgtgt ttttaataat 300
gtcagctatg tgccccatcc tccttcatgc cctccctccc tttcctacca ctgctgagtg 360
gcctggaact tgtttaaagt gt 382

<210> 140
<211> 200
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(200)
<223> n = A,T,C or G

<400> 140
accaaanctt ctttctgttg tgttngattt tactataggg gtttngcttn ttctaaanat 60
acttttcatt taacancttt tgttaagtgt caggctgcac tttgctccat anaattattg 120
ttttcacatt tcaacttgta tgtgtttgtc tcttanagca ttggtgaaat cacatatttt 180
atattcagca taaaggagaa 200

<210> 141
<211> 335
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(335)
<223> n = A,T,C or G

<400> 141
actttatttt caaaacactc atatgttgca aaaaacacat agaaaaataa agtttggtgg 60
gggtgctgac taaacttcaa gtcacagact tttatgtgac agattggagc agggtttgtt 120
atgcatgtag agaacccaaa ctaatttatt aaacaggata gaaacaggct gtctgggtga 180
aatggttctg agaaccatcc aattcacctg tcagatgctg atanactagc tcttcagatg 240
tttttctacc agttcagaga tnggttaatg actanttcca atggggaaaa agcaagatgg 300
attcacaaac caagtaattt taaacaaaga cactt 335

<210> 142
<211> 459
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(459)
<223> n = A,T,C or G

<400> 142
accaggttaa tattgccaca tatatccttt ccaattgcgg gctaaacaga cgtgtattta 60
gggttgttta aagacaaccc agcttaatat caagagaaat tgtgaccttt catggagtat 120
ctgatggaga aaacactgag ttttgacaaa tcttatttta ttcagatagc agtctgatca 180
cacatggtcc aacaacactc aaataataaa tcaaatatna tcagatgtta aagattggtc 240
ttcaaacatc atagccaatg atgccccgct tgcctataat ctctccgaca taaaaccaca 300
tcaacacctc agtggccacc aaaccattca gcacagcttc cttaactgtg agctgtttga 360
agctaccagt ctgagcacta ttgactatnt ttttcangct ctgaatagct ctagggatct 420
cagcangggt gggaggaacc agctcaacct tggcgtant 459

<210> 143
<211> 140
<212> DNA
<213> Homo sapien

<400> 143
acatttcctt ccaccaagtc aggactcctg gcttctgtgg gagttcttat cacctgaggg 60
aaatccaaac agtctctcct agaaaggaat agtgtcacca accccaccca tctccctgag 120
accatccgac ttccctgtgt 140

<210> 144
<211> 164
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(164)
<223> n = A,T,C or G

<400> 144
acttcagtaa caacatacaa taacaacatt aagtgtatat tgccatcttt gtcattttct 60
atctatacca ctctcccttc tgaaaacaan aatcactanc caatcactta tacaaatttg 120
aggcaattaa tccatatttg ttttcaataa ggaaaaaaag atgt 164

<210> 145
<211> 303
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(303)
<223> n = A,T,C or G

<400> 145
acgtagacca tccaactttg tatttgtaat ggcaaacatc cagnagcaat tcctaaacaa 60
actggagggt atttataccc aattatccca ttcattaaca tgccctcctc ctcaggctat 120
gcaggacagc tatcataagt cggcccaggc atccagatac taccatttgt ataaacttca 180
gtaggggagt ccatccaagt gacaggtcta atcaaaggag gaaatggaac ataagcccag 240
tagtaaaatn ttgcttagct gaaacagcca caaaagactt accgccgtgg tgattaccat 300
caa 303

<210> 146
<211> 327
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(327)
<223> n = A,T,C or G

<400> 146
actgcagctc aattagaagt ggtctctgac tttcatcanc ttctccctgg gctccatgac 60
actggcctgg agtgactcat tgctctggtt ggttgagaga gctcctttgc caacaggcct 120
ccaagtcagg gctgggattt gtttcctttc cacattctag caacaatatg ctggccactt 180
cctgaacagg gagggtggga ggagccagca tggaacaagc tgccactttc taaagtagcc 240
agacttgccc ctgggcctgt cacacctact gatgaccttc tgtgcctgca ggatggaatg 300
taggggtgag ctgtgtgact ctatggt 327

<210> 147
<211> 173
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(173)
<223> n = A,T,C or G

<400> 147
acattgtttt tttgagataa agcattgana gagctctcct taacgtgaca caatggaagg 60
actggaacac atacccacat ctttgttctg agggataatt ttctgataaa gtcttgctgt 120
atattcaagc acatatgtta tatattattc agttccatgt ttatagccta gtt 173

<210> 148

<211> 477
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(477)
<223> n = A,T,C or G

<400> 148
acaaccactt tatctcatcg aatttttaac ccaaactcac tcactgtgcc tttctatcct 60
atgggatata ttatttgatg ctccatttca tcacacatat atgaataata cactcatact 120
gccctactac ctgctgcaat aatcacattc ccttcctgtc ctgaccctga agccattggg 180
gtggtcctag tggccatcag tccangcctg caccttgagc ccttgagctc cattgctcac 240
nccancccac ctcaccgacc ccatcctctt acacagctac ctccttgctc tctaacccca 300
tagattatnt ccaaattcag tcaattaagt tactattaac actctacccg acatgtccag 360
caccactggt aagccttctc cagccaacac acacacacac acacncacac acacacatat 420
ccaggcacag gctacctcat cttcacaatc acccctttaa ttaccatgct atggtgg 477

<210> 149
<211> 207
<212> DNA
<213> Homo sapien

<400> 149
acagttgtat tataatatca agaaataaac ttgcaatgag agcatttaag agggaagaac 60
taacgtattt tagagagcca aggaaggttt ctgtggggag tgggatgtaa ggtggggcct 120
gatgataaat aagagtcagc caggtaagtg ggtggtgtgg tatgggcaca gtgaagaaca 180
tttcaggcag agggaacagc agtgaaa 207

<210> 150
<211> 111
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(111)
<223> n = A,T,C or G

<400> 150
accttgattt cattgctgct ctgatggaaa cccaactatc taatttagct aaaacatggg 60
cacttaaatg tggtcagtgt ttggacttgt taactantgg catctttggg t 111

<210> 151
<211> 196
<212> DNA
<213> Homo sapien

<400> 151
agcgcggcag gtcatattga acattccaga tacctatcat tactcgatgc tgttgataac 60
agcaagatgg ctttgaactc agggtcacca ccagctattg gaccttacta tgaaaaccat 120
ggataccaac cggaaaaccc ctatcccgca cagcccactg tggtccccac tgtctacgag 180
gtgcatccgg ctcagt 196

<210> 152
<211> 132
<212> DNA

<213> Homo sapien

<400> 152
acagcacttt cacatgtaag aagggagaaa ttcctaaatg taggagaaag ataacagaac 60
cttccccttt tcatctagtg gtggaaacct gatgctttat gttgacagga atagaaccag 120
gagggagttt gt 132

<210> 153
<211> 285
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(285)
<223> n = A,T,C or G

<400> 153
acaanaccca nganaggcca ctggccgtgg tgtcatggcc tccaaacatg aaagtgtcag 60
cttctgctct tatgtcctca tctgacaact ctttaccatt tttatcctcg ctcagcagga 120
gcacatcaat aaagtccaaa gtcttggact tggccttggc ttggaggaag tcatcaacac 180
cctggctagt gagggtgcgg cgccgctcct ggatgacggc atctgtgaag tcgtgcacca 240
gtctgcaggc cctgtggaag cgccgtccac acggagtnag gaatt 285

<210> 154
<211> 333
<212> DNA
<213> Homo sapien

<400> 154
accacagtcc tgttgggcca gggcttcatg acccttttctg tgaaaagcca tattatcacc 60
accccaaatt tttccttaaa tatctttaac tgaaggggtc agcctcttga ctgcaaagac 120
cctaagccgg ttacacagct aactcccact ggccctgatt tgtgaaattg ctgctgcctg 180
attggcacag gagtcgaagg tgttcagctc ccctcctccg tggaacgaga ctctgatttg 240
agtttcacaa attctcgggc cacctcgtca ttgctcctct gaaataaaat ccggagaatg 300
gtcaggcctg tctcatccat atggatcttc cgg 333

<210> 155
<211> 308
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(308)
<223> n = A,T,C or G

<400> 155
actggaaata ataaaaccca catcacagtg ttgtgtcaaa gatcatcagg gcatggatgg 60
gaaagtgctt tgggaactgt aaagtgccta acacatgatc gatgattttt gttataatat 120
ttgaatcacg gtgcatacaa actctcctgc ctgctcctcc tgggccccag ccccagcccc 180
atcacagctc actgctctgt tcatccaggc ccagcatgta gtggctgatt cttcttggct 240
gcttttagcc tccanaagtt tctctgaagc caaccaaacc tctangtgta aggcatgctg 300
gccctggt 308

<210> 156
<211> 295
<212> DNA

<213> Homo sapien

<400> 156
accttgctcg gtgcttggaa catattagga actcaaaata tgagatgata acagtgccta 60
ttattgatta ctgagagaac tgttagacat ttagttgaag attttctaca caggaactga 120
gaataggaga ttatgtttgg ccctcatatt ctctcctatc ctccttgcct cattctatgt 180
ctaatatatt ctcaatcaaa taaggttagc ataatcagga aatcgaccaa ataccaatat 240
aaaaccagat gtctatcctt aagattttca aatagaaaac aaattaacag actat 295

<210> 157
<211> 126
<212> DNA
<213> Homo sapien

<400> 157
acaagtttaa atagtgctgt cactgtgcat gtgctgaaat gtgaaatcca ccacatttct 60
gaagagcaaa acaaattctg tcatgtaatc tctatcttgg gtcgtgggta tatctgtccc 120
cttagt 126

<210> 158
<211> 442
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(442)
<223> n = A,T,C or G

<400> 158
acccactggt cttggaaaca cccatcctta atacgatgat ttttctgtcg tgtgaaaatg 60
aanccagcag gctgccccta gtcagtcctt ccttccagag aaaaagagat ttgagaaagt 120
gcctgggtaa ttcaccatta atttcctccc ccaaactctc tgagtcttcc cttaatattt 180
ctggtggttc tgaccaaagc aggtcatggt ttgttgagca tttgggatcc cagtgaagta 240
natgtttgta gccttgcata cttagccctt cccacgcaca aacggagtgg cagagtggtg 300
ccaaccctgt tttcccagtc cacgtagaca gattcacagt gcggaattct ggaagctgga 360
nacagacggg ctctttgcag agccgggact ctgagangga catgagggcc tctgcctctg 420
tgttcattct ctgatgtcct gt 442

<210> 159
<211> 498
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(498)
<223> n = A,T,C or G

<400> 159
acttccaggt aacgttgttg tttccgttga gcctgaactg atgggtgacg ttgtaggttc 60
tccaacaaga actgaggttg cagagcgggt agggaagagt gctgttccag ttgcacctgg 120
gctgctgtgg actgttgttg attcctcact acggcccaag gttgtggaac tggcanaaag 180
gtgtgtttgtt ggganttgagc tcgggcggct gtggtaggtt gtgggctctt caacaggggc 240
tgctgtggtg ccgggangtg aangtgttgt gtcacttgag cttggccagc tctggaaagt 300
antanattct tcctgaaggc cagcgcttgt ggagctggca ngggtcantg ttgtgtgtaa 360
cgaaccagtg ctgctgtggg tgggtgtana tcctccacaa agcctgaagt tatggtgtcn 420
tcaggtaana atgtggtttc agtgtccctg ggcngctgtg gaaggttgta nattgtcacc 480

aagggaataa gctgtggt 498

<210> 160
<211> 380
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(380)
<223> n = A,T,C or G

<400> 160
acctgcatcc agcttccctg ccaaactcac aaggagacat caacctctag acagggaaac 60
agcttcagga tacttccagg agacagagcc accagcagca aaacaaatat tcccatgcct 120
ggagcatggc atagaggaag ctganaaatg tggggtctga ggaagccatt tgagtctggc 180
cactagacat ctcatcagcc acttgtgtga agagatgccc catgacccca gatgcctctc 240
ccacccttac ctccatctca cacacttgag ctttccactc tgtataattc taacatcctg 300
gagaaaaatg gcagtttgac cgaacctgtt cacaacggta gaggctgatt tctaacgaaa 360
cttgtagaat gaagcctgga 380

<210> 161
<211> 114
<212> DNA
<213> Homo sapien

<400> 161
actccacatc ccctctgagc aggcggttgt cgttcaaggt gtatttggcc ttgcctgtca 60
cactgtccac tggcccctta tccacttggt gcttaatccc tcgaaagagc atgt 114

<210> 162
<211> 177
<212> DNA
<213> Homo sapien

<400> 162
actttctgaa tcgaatcaaa tgatacttag tgtagtttta atatcctcat atatatcaaa 60
gttttactac tctgataatt ttgtaaacca ggtaaccaga acatccagtc atacagcttt 120
tggtgatata taacttggca ataacccagt ctggtgatac ataaaactac tcactgt 177

<210> 163
<211> 137
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(137)
<223> n = A,T,C or G

<400> 163
catttataca gacaggcgtg aagacattca cgacaaaaac gcgaaattct atcccgtgac 60
canagaaggc agctacggct actcctacat cctggcgtgg gtggccttcg cctgcacctt 120
catcagcggc atgatgt 137

<210> 164
<211> 469
<212> DNA

<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(469)
<223> n = A,T,C or G

<400> 164
cttatcacaa tgaatgttct cctgggcagc gttgtgatct ttgccacctt cgtgacttta 60
tgcaatgcat catgctattt catacctaat gagggagttc caggagattc aaccaggaaa 120
tgcatggatc tcaaaggaaa caaacaccca ataaactcgg agtggcagac tgacaactgt 180
gagacatgca cttgctacga aacagaaatt tcatgttgca cccttgtttc tacacctgtg 240
ggttatgaca aagacaactg ccaaagaatc ttcaagaagg aggactgcaa gtatatcgtg 300
gtggagaaga aggacccaaa aaagacctgt tctgtcagtg aatggataat ctaatgtgct 360
tctagtaggc acagggctcc caggccaggc ctcattctcc tctggcctct aatagtcaat 420
gattgtgtag ccatgcctat cagtaaaaag atntttgagc aaacacttt 469

<210> 165
<211> 195
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(195)
<223> n = A,T,C or G

<400> 165
acagtttttt atanatatcg acattgccgg cacttgtgtt cagtttcata aagctggtgg 60
atccgctgtc atccactatt ccttggctag agtaaaaatt attcttatag cccatgtccc 120
tgcaggccgc ccgcccgtag ttctcgttcc agtcgtcttg gcacacaggg tgccaggact 180
tcctctgaga tgagt 195

<210> 166
<211> 383
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(383)
<223> n = A,T,C or G

<400> 166
acatcttagt agtgtggcac atcaggggge catcagggtc acagtcactc atagcctcgc 60
cgaggtcgga gtccacacca ccggtgtagg tgtgctcaat cttgggcttg gcgcccacct 120
ttggagaagg gatatgctgc acacacatgt ccacaaagcc tgtgaactcg ccaaagaatt 180
tttgcagacc agcctgagca aggggcggat gttcagcttc agctcctcct tcgtcaggtg 240
gatgccaacc tcgtctangg tccgtgggaa gctggtgtcc acntcaccta caacctgggc 300
gangatctta taaagaggct ccnagataaa ctccacgaaa cttctctggg agctgctagt 360
nggggccttt ttggtgaact ttc 383

<210> 167
<211> 247
<212> DNA
<213> Homo sapien

<220>

<221> misc_feature
<222> (1)...(247)
<223> n = A,T,C or G

<400> 167
acagagccag accttggcca taaatgaanc agagattaag actaaacccc aagtcganat 60
tggagcagaa actggagcaa gaagtgggcc tggggctgaa gtagagacca aggccactgc 120
tatanccata cacagagcca actctcaggc caaggcnatg gttggggcag anccagagac 180
tcaatctgan tccaaagtgg tggctggaac actggtcatg acanaggcag tgactctgac 240
tgangtc 247

<210> 168
<211> 273
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(273)
<223> n = A,T,C or G

<400> 168
acttctaagt tttctagaag tggaaggatt gtantcatcc tgaaaatggg tttacttcaa 60
aatccctcan ccttgttctt cacnactgtc tatactgana gtgtcatgtt tccacaaagg 120
gctgacacct gagcctgnat tttcactcat ccctgagaag ccctttccag tagggtgggc 180
aattcccaac ttccttgcca caagcttccc aggctttctc ccctggaaaa ctccagcttg 240
agtcccagat acactcatgg gctgccctgg gca 273

<210> 169
<211> 431
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(431)
<223> n = A,T,C or G

<400> 169
acagccttgg cttccccaaa ctccacagtc tcagtgcaga aagatcatct tccagcagtc 60
agctcagacc agggtcaaag gatgtgacat caacagtttc tggtttcaga acaggttcta 120
ctactgtcaa atgacccccc atacttcctc aaaggctgtg gtaagttttg cacaggtgag 180
ggcagcagaa agggggtant tactgatgga caccatcttc tctgtatact ccacactgac 240
cttgccatgg gcaaaggccc ctaccacaaa aacaatagga tcactgctgg gcaccagctc 300
acgcacatca ctgacaaccg ggatggaaaa agaantgcca actttcatac atccaactgg 360
aaagtgatct gatactggat tcttaattac cttcaaaagc ttctgggggc catcagctgc 420
tcgaacactg a 431

<210> 170
<211> 266
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(266)
<223> n = A,T,C or G

<400> 170
acctgtgggc tgggctgtta tgcctgtgcc ggctgctgaa agggagttca gaggtggagc 60
tcaaggagct ctgcaggcat tttgccaanc ctctccanag canagggagc aacctacact 120
ccccgctaga aagacaccag attggagtcc tgggaggggg agttggggtg ggcatttgat 180
gtatacttgt cacctgaatg aangagccag agaggaanga gacgaanatg anattggcct 240
tcaaagctag gggtctggca ggtgga 266

<210> 171
<211> 1248
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1248)
<223> n = A,T,C or G

<400> 171
ggcagccaaa tcataaacgg cgaggactgc agcccgcact cgcagccctg gcaggcggca 60
ctggtcatgg aaaacgaatt gttctgctcg ggcgtcctgg tgcatccgca gtgggtgctg 120
tcagccgcac actgtttcca gaagtgagtg cagagctcct acaccatcgg gctgggcctg 180
cacagtcttg aggccgacca agagccaggg agccagatgg tggaggccag cctctccgta 240
cggcacccag agtacaacag acccttgctc gctaacgacc tcatgctcat caagttggac 300
gaatccgtgt ccgagtctga caccatccgg agcatcagca ttgcttcgca gtgccctacc 360
gcggggaact cttgcctcgt ttctggctgg ggtctgctgg cgaacggcag aatgcctacc 420
gtgctgcagt gcgtgaacgt gtcggtggtg tctgaggagg tctgcagtaa gctctatgac 480
ccgctgtacc accccagcat gttctgcgcc ggcggagggc aagaccagaa ggactcctgc 540
aacggtgact ctgggggcgc cctgatctgc aacgggtact tgcagggcct tgtgtctttc 600
ggaaaagccc cgtgtggcca agttggcgtg ccaggtgtct acaccaacct ctgcaaattc 660
actgagtgga tagagaaaac cgtccaggcc agttaactct ggggactggg aacccatgaa 720
attgaccccc aaatacatcc tgcggaagga attcaggaat atctgttccc agcccctcct 780
ccctcaggcc caggagtcca ggcccccagc ccctcctccc tcaaaccaag ggtacagatc 840
cccagcccct cctccctcag acccaggagt ccagaccccc cagcccctcc tccctcagac 900
ccaggagtcc agcccctcct ccctcagacc caggagtcca gacccccag ccctcctcc 960
ctcagaccca ggggtccagg cccccaaccc ctcctccctc agactcagag gtccaagccc 1020
ccaacccntc attccccaga cccagaggtc caggtcccag cccctcntcc ctcagaccca 1080
gcggtccaat gccacctaga ctntccctgt acacagtgcc cccttgtggc acgttgaccc 1140
aaccttacca gttggttttt catttttngt ccctttcccc tagatccaga aataaagttt 1200
aagagaagng caaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa 1248

<210> 172
<211> 159
<212> PRT
<213> Homo sapien

<220>
<221> VARIANT
<222> (1)...(159)
<223> Xaa = Any Amino Acid

<400> 172
Met Val Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Pro
1 5 10 15
Leu Leu Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser
20 25 30
Glu Ser Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr
35 40 45
Ala Gly Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Gly

```
    50                  55                  60
Arg Met Pro Thr Val Leu Gln Cys Val Asn Val Ser Val Val Ser Glu
65                  70                  75                  80
Glu Val Cys Ser Lys Leu Tyr Asp Pro Leu Tyr His Pro Ser Met Phe
                85                  90                  95
Cys Ala Gly Gly Gly Gln Xaa Gln Xaa Asp Ser Cys Asn Gly Asp Ser
            100                 105                 110
Gly Gly Pro Leu Ile Cys Asn Gly Tyr Leu Gln Gly Leu Val Ser Phe
        115                 120                 125
Gly Lys Ala Pro Cys Gly Gln Val Gly Val Pro Gly Val Tyr Thr Asn
    130                 135                 140
Leu Cys Lys Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Ala Ser
145                 150                 155
```

<210> 173
<211> 1265
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1265)
<223> n = A,T,C or G

<400> 173

```
ggcagcccgc actcgcagcc ctggcaggcg gcactggtca tggaaaacga attgttctgc      60
tcgggcgtcc tggtgcatcc gcagtgggtg ctgtcagccg cacactgttt ccagaactcc     120
tacaccatcg ggctgggcct gcacagtctt gaggccgacc aagagccagg gagccagatg     180
gtggaggcca gcctctccgt acggcaccca gagtacaaca gacccttgct cgctaacgac     240
ctcatgctca tcaagttgga cgaatccgtg tccgagtctg acaccatccg gagcatcagc     300
attgcttcgc agtgccctac cgcggggaac tcttgcctcg tttctggctg gggtctgctg     360
gcgaacggtg agctcacggg tgtgtgtctg ccctcttcaa ggaggtcctc tgcccagtcg     420
cgggggctga cccagagctc tgcgtcccag gcagaatgcc taccgtgctg cagtgcgtga     480
acgtgtcggt ggtgtctgag gaggtctgca gtaagctcta tgacccgctg taccacccca     540
gcatgttctg cgccggcgga gggcaagacc agaaggactc ctgcaacggt gactctgggg     600
ggcccctgat ctgcaacggg tacttgcagg gccttgtgtc tttcggaaaa gccccgtgtg     660
gccaagttgg cgtgccaggt gtctacacca acctctgcaa attcactgag tggatagaga     720
aaaccgtcca ggccagttaa ctctggggac tgggaaccca tgaaattgac ccccaaatac     780
atcctgcgga aggaattcag gaatatctgt tcccagcccc tcctccctca ggcccaggag     840
tccaggcccc cagcccctcc tccctcaaac caagggtaca gatccccagc ccctcctccc     900
tcagacccag gagtccagac cccccagccc ctcctccctc agacccagga gtccagcccc     960
tcctccntca gacccaggag tccagacccc ccagcccctc ctccctcaga cccaggggtt    1020
gaggccccca acccctcctc cttcagagtc agaggtccaa gcccccaacc cctcgttccc    1080
cagacccaga ggtnnaggtc ccagcccctc ttccntcaga cccagnggtc caatgccacc    1140
tagattttcc ctgnacacag tgccccctg tggnangttg acccaacctt accagttggt    1200
ttttcatttt tngtcccttt cccctagatc cagaaataaa gtttaagaga ngngcaaaaa    1260
aaaaa                                                               1265
```

<210> 174
<211> 1459
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1459)
<223> n = A,T,C or G

<400> 174

ggtcagccgc acactgtttc cagaagtgag tgcagagctc ctacaccatc gggctgggcc 60
tgcacagtct tgaggccgac caagagccag ggagccagat ggtggaggcc agcctctccg 120
tacggcaccc agagtacaac agacccttgc tcgctaacga cctcatgctc atcaagttgg 180
acgaatccgt gtccgagtct gacaccatcc ggagcatcag cattgcttcg cagtgcccta 240
ccgcggggaa ctcttgcctc gtttctggct ggggtctgct ggcgaacggt gagctcacgg 300
gtgtgtgtct gccctcttca aggaggtcct ctgcccagtc gcgggggctg acccagagct 360
ctgcgtccca ggcagaatgc ctaccgtgct gcagtgcgtg aacgtgtcgg tggtgtctga 420
ngaggtctgc antaagctct atgacccgct gtaccacccc ancatgttct gcgccggcgg 480
agggcaagac cagaaggact cctgcaacgt gagagagggg aaaggggagg gcaggcgact 540
cagggaaggg tggagaaggg ggagacagag acacacaggg ccgcatggcg agatgcagag 600
atggagagac acacagggag acagtgacaa ctagagagag aaactgagag aaacagagaa 660
ataaacacag gaataaagag aagcaaagga agagagaaac agaaacagac atggggaggc 720
agaaacacac acacatagaa atgcagttga ccttccaaca gcatggggcc tgagggcggt 780
gacctccacc caatagaaaa tcctcttata acttttgact ccccaaaaac ctgactagaa 840
atagcctact gttgacgggg agccttacca ataacataaa tagtcgattt atgcatacgt 900
tttatgcatt catgatatac ctttgttgga atttttgat atttctaagc tacacagttc 960
gtctgtgaat ttttttaaat tgttgcaact ctcctaaaat ttttctgatg tgtttattga 1020
aaaaatccaa gtataagtgg acttgtgcat tcaaaccagg gttgttcaag ggtcaactgt 1080
gtacccagag ggaaacagtg acacagattc atagaggtga aacacgaaga gaaacaggaa 1140
aaatcaagac tctacaaaga ggctgggcag ggtggctcat gcctgtaatc ccagcacttt 1200
gggaggcgag gcaggcagat cacttgaggt aaggagttca agaccagcct ggccaaaatg 1260
gtgaaatcct gtctgtacta aaaatacaaa agttagctgg atatggtggc aggcgcctgt 1320
aatcccagct acttgggagg ctgaggcagg agaattgctt gaatatggga ggcagaggtt 1380
gaagtgagtt gagatcacac cactatactc cagctggggc aacagagtaa gactctgtct 1440
caaaaaaaaa aaaaaaaaa 1459

<210> 175
<211> 1167
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1167)
<223> n = A,T,C or G

<400> 175

gcgcagccct ggcaggcggc actggtcatg gaaaacgaat tgttctgctc gggcgtcctg 60
gtgcatccgc agtgggtgct gtcagccgca cactgtttcc agaactccta caccatcggg 120
ctgggcctgc acagtcttga ggccgaccaa gagccaggga gccagatggt ggaggccagc 180
ctctccgtac ggcacccaga gtacaacaga ctcttgctcg ctaacgacct catgctcatc 240
aagttggacg aatccgtgtc cgagtctgac accatccgga gcatcagcat tgcttcgcag 300
tgccctaccg cggggaactc ttgcctcgtn tctggctggg gtctgctggc gaacggcaga 360
atgcctaccg tgctgcactg cgtgaacgtg tcggtggtgt ctgaggangt ctgcagtaag 420
ctctatgacc cgctgtacca ccccagcatg ttctgcgccg gcggagggca agaccagaag 480
gactcctgca acggtgactc tgggggggccc ctgatctgca acgggtactt gcagggcctt 540
gtgtctttcg gaaaagcccc gtgtggccaa cttggcgtgc caggtgtcta caccaacctc 600
tgcaaattca ctgagtggat agagaaaacc gtccagncca gttaactctg gggactggga 660
acccatgaaa ttgaccccca aatacatcct gcggaangaa ttcaggaata tctgttccca 720
gcccctcctc cctcaggccc aggagtccag gcccccagcc cctcctccct caaaccaagg 780
gtacagatcc ccagcccctc ctccctcaga cccaggagtc cagacccccc agcccctcnt 840
ccntcagacc caggagtcca gcccctcctc cntcagacgc aggagtccag acccccccagc 900
ccntcntccg tcagacccag gggtgcaggc ccccaacccc tcntccntca gagtcagagg 960
tccaagcccc caacccctcg ttccccagac ccagaggtnc aggtcccagc ccctcctccc 1020
tcagacccag cggtccaatg ccacctagan tntccctgta cacagtgccc ccttgtggca 1080
ngttgaccca accttaccag ttggtttttc atttttgtc cctttcccct agatccagaa 1140
ataaagtnta agagaagcgc aaaaaaa 1167

<210> 176
<211> 205
<212> PRT
<213> Homo sapien

<220>
<221> VARIANT
<222> (1)...(205)
<223> Xaa = Any Amino Acid

<400> 176

```
Met Glu Asn Glu Leu Phe Cys Ser Gly Val Leu Val His Pro Gln Trp
 1               5                  10                  15
Val Leu Ser Ala Ala His Cys Phe Gln Asn Ser Tyr Thr Ile Gly Leu
            20                  25                  30
Gly Leu His Ser Leu Glu Ala Asp Gln Glu Pro Gly Ser Gln Met Val
        35                  40                  45
Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Leu Leu Leu
    50                  55                  60
Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser Glu Ser
65                  70                  75                  80
Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala Gly
                85                  90                  95
Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Gly Arg Met
            100                 105                 110
Pro Thr Val Leu His Cys Val Asn Val Ser Val Val Ser Glu Xaa Val
        115                 120                 125
Cys Ser Lys Leu Tyr Asp Pro Leu Tyr His Pro Ser Met Phe Cys Ala
    130                 135                 140
Gly Gly Gly Gln Asp Gln Lys Asp Ser Cys Asn Gly Asp Ser Gly Gly
145                 150                 155                 160
Pro Leu Ile Cys Asn Gly Tyr Leu Gln Gly Leu Val Ser Phe Gly Lys
                165                 170                 175
Ala Pro Cys Gly Gln Leu Gly Val Pro Gly Val Tyr Thr Asn Leu Cys
            180                 185                 190
Lys Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Xaa Ser
        195                 200                 205
```

<210> 177
<211> 1119
<212> DNA
<213> Homo sapien

<400> 177

```
gcgcactcgc agccctggca ggcggcactg gtcatggaaa acgaattgtt ctgctcgggc      60
gtcctggtgc atccgcagtg ggtgctgtca gccgcacact gtttccagaa ctcctacacc     120
atcgggctgg gcctgcacag tcttgaggcc gaccaagagc cagggagcca gatggtggag     180
gccagcctct ccgtacggca cccagagtac aacagaccct tgctcgctaa cgacctcatg     240
ctcatcaagt tggacgaatc cgtgtccgag tctgacacca tccggagcat cagcattgct     300
tcgcagtgcc ctaccgcggg gaactcttgc ctcgtttctg gctggggtct gctggcgaac     360
gatgctgtga ttgccatcca gtcccagact gtgggaggct gggagtgtga gaagctttcc     420
caaccctggc agggttgtac catttcggca acttccagtg caaggacgtc ctgctgcatc     480
ctcactgggt gctcactact gctcactgca tcacccggaa cactgtgatc aactagccag     540
caccatagtt ctccgaagtc agactatcat gattactgtg ttgactgtgc tgtctattgt     600
actaaccatg ccgatgttta ggtgaaatta gcgtcacttg gcctcaacca tcttggtatc     660
cagttatcct cactgaattg agatttcctg cttcagtgtc agccattccc acataatttc     720
tgacctacag aggtgaggga tcatatagct cttcaaggat gctggtactc ccctcacaaa     780
```

ttcatttctc ctgttgtagt gaaaggtgcg ccctctggag cctcccaggg tgggtgtgca 840
ggtcacaatg atgaatgtat gatcgtgttc ccattaccca aagcctttaa atccctcatg 900
ctcagtacac cagggcaggt ctagcatttc ttcatttagt gtatgctgtc cattcatgca 960
accacctcag gactcctgga ttctctgcct agttgagctc ctgcatgctg cctccttggg 1020
gaggtgaggg agagggccca tggttcaatg ggatctgtgc agttgtaaca cattaggtgc 1080
ttaataaaca gaagctgtga tgttaaaaaa aaaaaaaaa 1119

<210> 178
<211> 164
<212> PRT
<213> Homo sapien

<220>
<221> VARIANT
<222> (1)...(164)
<223> Xaa = Any Amino Acid

<400> 178

Met Glu Asn Glu Leu Phe Cys Ser Gly Val Leu Val His Pro Gln Trp
1 5 10 15
Val Leu Ser Ala Ala His Cys Phe Gln Asn Ser Tyr Thr Ile Gly Leu
20 25 30
Gly Leu His Ser Leu Glu Ala Asp Gln Glu Pro Gly Ser Gln Met Val
35 40 45
Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Pro Leu Leu
50 55 60
Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser Glu Ser
65 70 75 80
Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala Gly
85 90 95
Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Asp Ala Val
100 105 110
Ile Ala Ile Gln Ser Xaa Thr Val Gly Gly Trp Glu Cys Glu Lys Leu
115 120 125
Ser Gln Pro Trp Gln Gly Cys Thr Ile Ser Ala Thr Ser Ser Ala Arg
130 135 140
Thr Ser Cys Cys Ile Leu Thr Gly Cys Ser Leu Leu Leu Thr Ala Ser
145 150 155 160
Pro Gly Thr Leu

<210> 179
<211> 250
<212> DNA
<213> Homo sapien

<400> 179
ctggagtgcc ttggtgtttc aagcccctgc aggaagcaga atgcaccttc tgaggcacct 60
ccagctgccc ccggccgggg gatgcgaggc tcggagcacc cttgcccggc tgtgattgct 120
gccaggcact gttcatctca gcttttctgt ccctttgctc ccggcaagcg cttctgctga 180
aagttcatat ctggagcctg atgtcttaac gaataaaggt cccatgctcc acccgaaaaa 240
aaaaaaaaaa 250

<210> 180
<211> 202
<212> DNA
<213> Homo sapien

<400> 180
actagtccag tgtggtggaa ttccattgtg ttgggcccaa cacaatggct acctttaaca 60
tcacccagac cccgcccctg cccgtgcccc acgctgctgc taacgacagt atgatgctta 120
ctctgctact cggaaactat ttttatgtaa ttaatgtatg ctttcttgtt tataaatgcc 180
tgatttaaaa aaaaaaaaaa aa 202

<210> 181
<211> 558
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(558)
<223> n = A,T,C or G

<400> 181
tccytttgkt naggtttkkg agacamcccк agacctwaan ctgtgtcaca gacttcyngg 60
aatgtttagg cagtgctagt aatttcytcg taatgattct gttattactt tcctnattct 120
ttattcctct ttcttctgaa gattaatgaa gttgaaaatt gaggtggata aatacaaaaa 180
ggtagtgtga tagtataagt atctaagtgc agatgaaagt gtgttatata tatccattca 240
aaattatgca agttagtaat tactcagggt taactaaatt actttaatat gctgttgaac 300
ctactctgtt ccttggctag aaaaaattat aaacaggact ttgttagttt gggaagccaa 360
attgataata ttctatgttc taaaagttgg gctatacata aattattaag aaatatggaw 420
ttttattccc aggaatatgg kgttcatttt atgaatatta cscrggatag awgtwtgagt 480
aaaaycagtt ttggtwaata ygtwaatatg tcmtaaataa acaakgcttt gacttatttc 540
caaaaaaaaa aaaaaaaa 558

<210> 182
<211> 479
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(479)
<223> n = A,T,C or G

<400> 182
acagggwttk grggatgcta agsccccrga rwtygtttga tccaaccctg gcttwttttc 60
agaggggaaa atggggccta gaagttacag mscatytagy tggtgcgmtg gcacccctgg 120
cstcacacag astcccgagt agctgggact acaggcacac agtcactgaa gcaggccctg 180
ttwgcaattc acgttgccac ctccaactta aacattcttc atatgtgatg tccttagtca 240
ctaaggttaa actttcccac ccagaaaagg caacttagat aaaatcttag agtactttca 300
tactmttcta agtcctcttc cagcctcact kkgagtcctm cytgggggtt gataggaant 360
ntctcttggc tttctcaata aartctctat ycatctcatg tttaatttgg tacgcatara 420
awtgstgara aaattaaaat gttctggtty mactttaaaa araaaaaaaa aaaaaaaaa 479

<210> 183
<211> 384
<212> DNA
<213> Homo sapien

<400> 183
aggcgggagc agaagctaaa gccaaagccc aagaagagtg gcagtgccag cactggtgcc 60
agtaccagta ccaataacag tgccagtgcc agtgccagca ccagtggtgg cttcagtgct 120
ggtgccagcc tgaccgccac tctcacattt gggctcttcg ctggccttgg tggagctggt 180
gccagcacca gtggcagctc tggtgcctgt ggtttctcct acaagtgaga ttttagatat 240

tgttaatcct gccagtcttt ctcttcaagc cagggtgcat cctcagaaac ctactcaaca 300
cagcactcta ggcagccact atcaatcaat tgaagttgac actctgcatt aratctattt 360
gccatttcaa aaaaaaaaaa aaaa 384

<210> 184
<211> 496
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(496)
<223> n = A,T,C or G

<400> 184
accgaattgg gaccgctggc ttataagcga tcatgtyynt ccrgtatkac ctcaacgagc 60
agggagatcg agtctatacg ctgaagaaat ttgacccgat gggacaacag acctgctcag 120
cccatcctgc tcggttctcc ccagatgaca aatactctsg acaccgaatc accatcaaga 180
aacgcttcaa ggtgctcatg acccagcaac cgcgccctgt cctctgaggg tcccttaaac 240
tgatgtcttt tctgccacct gttacccctc ggagactccg taaccaaact cttcggactg 300
tgagccctga tgccttttg ccagccatac tctttggcat ccagtctctc gtggcgattg 360
attatgcttg tgtgaggcaa tcatggtggc atcacccata aagggaacac atttgacttt 420
tttttctcat attttaaatt actacmagaw tattwmagaw waaatgawtt gaaaaactst 480
taaaaaaaaa aaaaaa 496

<210> 185
<211> 384
<212> DNA
<213> Homo sapien

<400> 185
gctggtagcc tatggcgkgg cccacggagg ggctcctgag gccacggrac agtgacttcc 60
caagtatcyt gcgcsgcgtc ttctaccgtc cctacctgca gatcttcggg cagattcccc 120
aggaggacat ggacgtggcc ctcatggagc acagcaactg ytcgtcggag cccggcttct 180
gggcacaccc tcctggggcc caggcgggca cctgcgtctc ccagtatgcc aactggctgg 240
tggtgctgct cctcgtcatc ttcctgctcg tggccaacat cctgctggtc aacttgctca 300
ttgccatgtt cagttacaca ttcggcaaag tacagggcaa cagcgatctc tactgggaag 360
gcgcagcgtt accgcctcat ccgg 384

<210> 186
<211> 577
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(577)
<223> n = A,T,C or G

<400> 186
gagttagctc ctccacaacc ttgatgaggt cgtctgcagt ggcctctcgc ttcataccgc 60
tnccatcgtc atactgtagg tttgccacca cytcctggca tcttggggcg gcntaatatt 120
ccaggaaact ctcaatcaag tcaccgtcga tgaaacctgt gggctggttc tgtcttccgc 180
tcggtgtgaa aggatctccc agaaggagtg ctcgatcttc cccacacttt tgatgacttt 240
attgagtcga ttctgcatgt ccagcaggag gttgtaccag ctctctgaca gtgaggtcac 300
cagccctatc atgccgttga mcgtgccgaa garcaccgag ccttgtgtgg gggkkgaagt 360
ctcacccaga ttctgcatta ccagagagcc gtggcaaaag acattgacaa actcgcccag 420
gtggaaaaag amcamctcct ggargtgctn gccgctcctc gtcmgttggt ggcagcgctw 480

tccttttgac acacaaacaa gttaaaggca ttttcagccc ccagaaantt gtcatcatcc 540
aagatntcgc acagcactna tccagttggg attaaat 577

<210> 187
<211> 534
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(534)
<223> n = A,T,C or G

<400> 187
aacatcttcc tgtataatgc tgtgtaatat cgatccgatn ttgtctgstg agaatycatw 60
actkggaaaa gmaacattaa agcctggaca ctggtattaa aattcacaat atgcaacact 120
ttaaacagtg tgtcaatctg ctcccyynac tttgtcatca ccagtctggg aakaagggta 180
tgccctattc acacctgtta aaagggcgct aagcattttt gattcaacat ctttttttt 240
gacacaagtc cgaaaaaagc aaaagtaaac agttatyaat ttgttagcca attcactttc 300
ttcatgggac agagccatyt gatttaaaaa gcaaattgca taatattgag cttygggagc 360
tgatatttga gcggaagagt agcctttcta cttcaccaga cacaactccc tttcatattg 420
ggatgttnac naaagtwatg tctctwacag atgggatgct tttgtggcaa ttctgttctg 480
aggatctccc agtttattta ccacttgcac aagaaggcgt tttcttcctc aggc 534

<210> 188
<211> 761
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(761)
<223> n = A,T,C or G

<400> 188
agaaaccagt atctctnaaa acaacctctc ataccttgtg gacctaattt tgtgtgcgtg 60
tgtgtgtgcg cgcatattat atagacaggc acatcttttt tacttttgta aaagcttatg 120
cctctttggt atctatatct gtgaaagttt taatgatctg ccataatgtc ttggggacct 180
ttgtcttctg tgtaaatggt actagagaaa acacctatnt tatgagtcaa tctagttngt 240
tttattcgac atgaaggaaa tttccagatn acaacactna caaactctcc ctkgackarg 300
ggggacaaag aaaagcaaaa ctgamcataa raaacaatwa cctggtgaga arttgcataa 360
acagaaatwr ggtagtatat tgaarnacag catcattaaa rmgttwtktt wttctccctt 420
gcaaaaaaca tgtacngact tcccgttgag taatgccaag ttgtttttt tatnataaaa 480
cttgcccttc attacatgtt tnaaagtggt gtggtgggcc aaaatattga aatgatggaa 540
ctgactgata aagctgtaca aataagcagt gtgcctaaca agcaacacag taatgttgac 600
atgcttaatt cacaaatgct aatttcatta taaatgtttg ctaaaataca ctttgaacta 660
ttttttctgtn ttcccagagc tgagatntta gattttatgt agtatnaagt gaaaaantac 720
gaaaataata acattgaaga aaaananaaa aaanaaaaaa a 761

<210> 189
<211> 482
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(482)
<223> n = A,T,C or G

<400> 189
tttttttttt tttgccgatn ctactatttt attgcaggan gtgggggtgt atgcaccgca 60
caccggggct atnagaagca agaaggaagg agggagggca cagccccttg ctgagcaaca 120
aagccgcctg ctgccttctc tgtctgtctc ctggtgcagg cacatgggga gaccttcccc 180
aaggcagggg ccaccagtcc aggggtggga atacaggggg tgggangtgt gcataagaag 240
tgataggcac aggccacccg gtacagaccc ctcggctcct gacaggtnga tttcgaccag 300
gtcattgtgc cctgcccagg cacagcgtan atctggaaaa gacagaatgc tttccttttc 360
aaatttggct ngtcatngaa ngggcanttt tccaanttng gctnggtctt ggtacncttg 420
gttcggccca gctccncgtc caaaaantat tcacccnnct ccnaattgct tgcnggnccc 480
cc 482

<210> 190
<211> 471
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(471)
<223> n = A,T,C or G

<400> 190
tttttttttt ttttaaaaca gtttttcaca acaaaattta ttagaagaat agtggttttg 60
aaaactctcg catccagtga gaactaccat acaccacatt acagctngga atgtnctcca 120
aatgtctggt caaatgatac aatggaacca ttcaatctta cacatgcacg aaagaacaag 180
cgcttttgac atacaatgca caaaaaaaaa agggggggggg gaccacatgg attaaaattt 240
taagtactca tcacatacat taagacacag ttctagtcca gtcnaaaatc agaactgcnt 300
tgaaaaattt catgtatgca atccaaccaa agaacttnat tggtgatcat gantnctcta 360
ctacatcnac cttgatcatt gccaggaacn aaaagttnaa ancacncngt acaaaaanaa 420
tctgtaattn anttcaacct ccgtacngaa aaatnttnnt tatacactcc c 471

<210> 191
<211> 402
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(402)
<223> n = A,T,C or G

<400> 191
gagggattga aggtctgttc tastgtcggm ctgttcagcc accaactcta acaagttgct 60
gtcttccact cactgtctgt aagcttttta acccagacwg tatcttcata aatagaacaa 120
attcttcacc agtcacatct tctaggacct ttttggattc agttagtata agctcttcca 180
cttcctttgt taagacttca tctggtaaag tcttaagttt tgtagaaagg aattyaattg 240
ctcgttctct aacaatgtcc tctccttgaa gtatttggct gaacaaccca cctaaagtcc 300
ctttgtgcat ccattttaaa tatacttaat agggcattgk tncactaggt taaattctgc 360
aagagtcatc tgtctgcaaa agttgcgtta gtatatctgc ca 402

<210> 192
<211> 601
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature

<222> (1)...(601)
<223> n = A,T,C or G

<400> 192
gagctcggat ccaataatct ttgtctgagg gcagcacaca tatncagtgc catggnaact 60
ggtctacccc acatgggagc agcatgccgt agntatataa ggtcattccc tgagtcagac 120
atgcytyttt gaytaccgtg tgccaagtgc tggtgattct yaacacacyt ccatcccgyt 180
cttttgtgga aaaactggca cttktctgga actagcarga catcacttac aaattcaccc 240
acgagacact tgaaaggtgt aacaaagcga ytcttgcatt gctttttgtc cctccggcac 300
cagttgtcaa tactaacccg ctggtttgcc tccatcacat ttgtgatctg tagctctgga 360
tacatctcct gacagtactg aagaacttct tcttttgttt caaaagcarc tcttggtgcc 420
tgttggatca ggttcccatt tcccagtcyg aatgttcaca tggcatattt wacttcccac 480
aaaacattgc gatttgaggc tcagcaacag caaatcctgt tccggcattg gctgcaagag 540
cctcgatgta gccggccagc gccaaggcag gcgccgtgag ccccaccagc agcagaagca 600
g 601

<210> 193
<211> 608
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(608)
<223> n = A,T,C or G

<400> 193
atacagccca natcccacca cgaagatgcg cttgttgact gagaacctga tgcggtcact 60
ggtcccgctg tagccccagc gactctccac ctgctggaag cggttgatgc tgcactcytt 120
cccaacgcag gcagmagcgg gsccggtcaa tgaactccay tcgtggcttg gggtkgacgg 180
tkaagtgcag gaagaggctg accacctcgc ggtccaccag gatgcccgac tgtgcgggac 240
ctgcagcgaa actcctcgat ggtcatgagc gggaagcgaa tgaggcccag ggccttgccc 300
agaaccttcc gcctgttctc tggcgtcacc tgcagctgct gccgctgaca ctcggcctcg 360
gaccagcgga caaacggcrt tgaacagccg cacctcacgg atgcccagtg tgtcgcgctc 420
caggammgsc accagcgtgt ccaggtcaat gtcggtgaag ccctccgcgg gtratggcgt 480
ctgcagtgtt tttgtcgatg ttctccaggc acaggctggc cagctgcggt tcatcgaaga 540
gtcgcgcctg cgtgagcagc atgaaggcgt tgtcggctcg cagttcttct tcaggaactc 600
cacgcaat 608

<210> 194
<211> 392
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(392)
<223> n = A,T,C or G

<400> 194
gaacggctgg accttgcctc gcattgtgct tgctggcagg gaataccttg gcaagcagyt 60
ccagtccgag cagccccaga ccgctgccgc ccgaagctaa gcctgcctct ggccttcccc 120
tccgcctcaa tgcagaacca gtagtgggag cactgtgttt agagttaaga gtgaacactg 180
tttgatttta cttgggaatt tcctctgtta tatagctttt cccaatgcta atttccaaac 240
aacaacaaca aaataacatg tttgcctgtt aagttgtata aaagtaggtg attctgtatt 300
taaagaaaat attactgtta catatactgc ttgcaatttc tgtatttatt gktnctstgg 360
aaataaatat agttattaaa ggttgtcant cc 392

<210> 195
<211> 502
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(502)
<223> n = A,T,C or G

<400> 195
ccsttkgagg ggtkaggkyc cagttyccga gtggaagaaa caggccagga gaagtgcgtg 60
ccgagctgag gcagatgttc ccacagtgac ccccagagcc stgggstata gtytctgacc 120
cctcncaagg aaagaccacs ttctggggac atgggctgga gggcaggacc tagaggcacc 180
aagggaaggc cccattccgg ggstgttccc cgaggaggaa gggaaggggc tctgtgtgcc 240
ccccasgagg aagaggccct gagtcctggg atcagacacc ccttcacgtg tatccccaca 300
caaatgcaag ctcaccaagg tcccctctca gtccccttcc stacaccctg amcggccact 360
gscscacacc cacccagagc acgccacccg ccatggggar tgtgctcaag gartcgcngg 420
gcarcgtgga catctngtcc cagaaggggg cagaatctcc aatagangga ctgarcmstt 480
gctnanaaaa aaaaanaaaa aa 502

<210> 196
<211> 665
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(665)
<223> n = A,T,C or G

<400> 196
ggttacttgg tttcattgcc accacttagt ggatgtcatt tagaaccatt ttgtctgctc 60
cctctggaag ccttgcgcag agcggacttt gtaattgttg gagaataact gctgaatttt 120
wagctgtttk gagttgatts gcaccactgc acccacaact tcaatatgaa aacyawttga 180
actwatttat tatcttgtga aaagtataac aatgaaaatt ttgttcatac tgtattkatc 240
aagtatgatg aaaagcaawa gatatatatt cttttattat gttaaattat gattgccatt 300
attaatcggc aaaatgtgga gtgtatgttc ttttcacagt aatatatgcc ttttgtaact 360
tcacttggtt attttattgt aaatgartta caaaattctt aatttaagar aatggtatgt 420
watatttatt tcattaattt ctttcctkgt ttacgtwaat tttgaaaaga wtgcatgatt 480
tcttgacaga aatcgatctt gatgctgtgg aagtagtttg acccacatcc ctatgagttt 540
ttcttagaat gtataaaggt tgtagcccat cnaacttcaa agaaaaaaat gaccacatac 600
tttgcaatca ggctgaaatg tggcatgctn ttctaattcc aactttataa actagcaaan 660
aagtg 665

<210> 197
<211> 492
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(492)
<223> n = A,T,C or G

<400> 197
ttttnttttt ttttttttgc aggaaggatt ccatttattg tggatgcatt ttcacaatat 60
atgtttattg gagcgatcca ttatcagtga aaagtatcaa gtgtttataa nattttagg 120

aaggcagatt cacagaacat gctngtcngc ttgcagtttt acctcgtana gatnacagag 180
aattatagtc naaccagtaa acnaggaatt tacttttcaa aagattaaat ccaaactgaa 240
caaaattcta ccctgaaact tactccatcc aaatattgga ataanagtca gcagtgatac 300
attctcttct gaactttaga ttttctagaa aaatatgtaa tagtgatcag gaagagctct 360
tgttcaaaag tacaacnaag caatgttccc ttaccatagg ccttaattca aactttgatc 420
catttcactc ccatcacggg agtcaatgct acctgggaca cttgtatttt gttcatnctg 480
ancntggctt aa 492

<210> 198
<211> 478
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(478)
<223> n = A,T,C or G

<400> 198
tttnttttgn atttcantct gtannaanta ttttcattat gtttattana aaaatatnaa 60
tgtntccacn acaaatcatn ttacntnagt aagaggccan ctacattgta caacatacac 120
tgagtatatt ttgaaaagga caagtttaaa gtanacncat attgccganc atancacatt 180
tatacatggc ttgattgata tttagcacag canaaactga gtgagttacc agaaanaaat 240
natatatgtc aatcngattt aagatacaaa acagatccta tggtacatan catcntgtag 300
gagttgtggc tttatgttta ctgaaagtca atgcagttcc tgtacaaaga gatggccgta 360
agcattctag tacctctact ccatggttaa gaatcgtaca cttatgttta catatgtnca 420
gggtaagaat tgtgttaagt naanttatgg agaggtccan gagaaaaatt tgatncaa 478

<210> 199
<211> 482
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(482)
<223> n = A,T,C or G

<400> 199
agtgacttgt cctccaacaa aacccccttga tcaagtttgt ggcactgaca atcagaccta 60
tgctagttcc tgtcatctat tcgctactaa atgcagactg gaggggacca aaaaggggca 120
tcaactccag ctggattatt ttggagcctg caaatctatt cctacttgta cggactttga 180
agtgattcag tttcctctac ggatgagaga ctggctcaag aatatcctca tgcagcttta 240
tgaagccnac tctgaacacg ctggttatct nagatgagaa ncagagaaat aaagtcnaga 300
aaatttacct ggangaaaag aggctttngg ctggggacca tcccattgaa ccttctctta 360
anggacttta agaanaaact accacatgtn tgtngtatcc tggtgccngg ccgtttantg 420
aacntngacn ncacccttnt ggaatanant cttgacngcn tcctgaactt gctcctctgc 480
ga 482

<210> 200
<211> 270
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(270)
<223> n = A,T,C or G

<400> 200
cggccgcaag tgcaactcca gctggggccg tgcggacgaa gattctgcca gcagttggtc 60
cgactgcgac gacggcggcg gcgacagtcg caggtgcagc gcgggcgcct ggggtcttgc 120
aaggctgagc tgacgccgca gaggtcgtgt cacgtcccac gaccttgacg ccgtcgggga 180
cagccggaac agagcccggt gaangcggga ggcctcgggg agcccctcgg gaagggcggc 240
ccgagagata cgcaggtgca ggtggccgcc 270

<210> 201
<211> 419
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(419)
<223> n = A,T,C or G

<400> 201
ttttttttnt ttttggaatc tactgcgagc acagcaggtc agcaacaagt ttattttgca 60
gctagcaagg taacagggta gggcatggtt acatgttcag gtcaacttcc tttgtcgtgg 120
ttgattggtt tgtctttatg ggggcggggt ggggtagggg aaancgaagc anaantaaca 180
tggagtgggt gcaccctccc tgtagaacct ggttacnaaa gcttggggca gttcacctgg 240
tctgtgaccg tcattttctt gacatcaatg ttattagaag tcaggatatc ttttagagag 300
tccactgtnt ctggagggag attagggttt cttgccaana tccaancaaa atccacntga 360
aaaagttgga tgatncangt acngaatacc ganggcatan ttctcatant cggtggcca 419

<210> 202
<211> 509
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(509)
<223> n = A,T,C or G

<400> 202
tttnttttnt ttttttnttt ttttttnttt ttttttnttt ttttttnttt ttttttnttt 60
tggcacttaa tccattttta tttcaaaatg tctacaaant ttnaatncnc cattatacng 120
gtnatttttnc aaaatctaaa nnttattcaa atntnagcca aantccttac ncaaatnnaa 180
tacncncaaa aatcaaaaat atacntntct ttcagcaaac ttngttacat aaattaaaaa 240
aatatatacg gctggtgttt tcaaagtaca attatcttaa cactgcaaac atntttnnaa 300
ggaactaaaa taaaaaaaaa cactnccgca aaggttaaag ggaacaacaa attcntttta 360
caacancnnc nattataaaa atcatatctc aaatcttagg ggaatatata cttcacacng 420
ggatcttaac ttttactnca ctttgtttat ttttttanaa ccattgtntt gggcccaaca 480
caatggnaat nccnccncnc tggactagt 509

<210> 203
<211> 583
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(583)
<223> n = A,T,C or G

<400> 203
ttttttttt ttttttttga ccccctctt ataaaaaaca agttaccatt ttatttact 60
tacacatatt tattttataa ttggtattag atattcaaaa ggcagctttt aaaatcaaac 120
taaatggaaa ctgccttaga tacataattc ttaggaatta gcttaaaatc tgcctaaagt 180
gaaaatcttc tctagctctt ttgactgtaa atttttgact cttgtaaaac atccaaattc 240
atttttcttg tctttaaaat tatctaatct ttccattttt tccctattcc aagtcaattt 300
gcttctctag cctcatttcc tagctcttat ctactattag taagtggctt ttttcctaaa 360
agggaaaaca ggaagagana atggcacaca aaacaaacat tttatattca tatttctacc 420
tacgttaata aaatagcatt ttgtgaagcc agctcaaaag aaggcttaga tccttttatg 480
tccattttag tcactaaacg atatcnaaag tgccagaatg caaaaggttt gtgaacattt 540
attcaaaagc taatataaga tatttcacat actcatcttt ctg 583

<210> 204
<211> 589
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(589)
<223> n = A,T,C or G

<400> 204
ttttttttnt ttttttttt ttttttnctc ttctttttt ttganaatga ggatcgagtt 60
tttcactctc tagatagggc atgaagaaaa ctcatctttc cagctttaaa ataacaatca 120
aatctcttat gctatatcat attttaagtt aaactaatga gtcactggct tatcttctcc 180
tgaaggaaat ctgttcattc ttctcattca tatagttata tcaagtacta ccttgcatat 240
tgagaggttt ttcttctcta tttacacata tatttccatg tgaatttgta tcaaaccttt 300
attttcatgc aaactagaaa ataatgtntt cttttgcata agagaagaga acaatatnag 360
cattacaaaa ctgctcaaat tgtttgttaa gnttatccat tataattagt tnggcaggag 420
ctaatacaaa tcacatttac ngacnagcaa taataaaact gaagtaccag ttaaatatcc 480
aaaataatta aaggaacatt tttagcctgg gtataattag ctaattcact ttacaagcat 540
ttattnagaa tgaattcaca tgttattatt ccntagccca acacaatgg 589

<210> 205
<211> 545
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(545)
<223> n = A,T,C or G

<400> 205
tttttnttttt ttttttcagt aataatcaga acaatattta tttttatatt taaaattcat 60
agaaaagtgc cttacattta ataaaagttt gtttctcaaa gtgatcagag gaattagata 120
tngtcttgaa caccaatatt aatttgagga aaatacacca aaatacatta agtaaattat 180
ttaagatcat agagcttgta agtgaaaaga taaaatttga cctcagaaac tctgagcatt 240
aaaaatccac tattagcaaa taaattacta tggacttctt gctttaattt tgtgatgaat 300
atggggtgtc actggtaaac caacacattc tgaaggatac attacttagt gatagattct 360
tatgtacttt gctanatnac gtggatatga gttgacaagt ttctctttct tcaatctttt 420
aaggggcnga ngaaatgagg aagaaaagaa aaggattacg catactgttc tttctatngg 480
aaggattaga tatgtttcct ttgccaatat taaaaaaata ataatgttta ctactagtga 540
aaccc 545

<210> 206
<211> 487

<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(487)
<223> n = A,T,C or G

<400> 206
tttttttttt ttttttagtc aagtttctna tttttattat aattaaagtc ttggtcattt 60
catttattag ctctgcaact tacatattta aattaaagaa acgttnttag acaactgtna 120
caatttataa atgtaaggtg ccattattga gtanatatat tcctccaaga gtggatgtgt 180
cccttctccc accaactaat gaancagcaa cattagttta attttattag tagatnatac 240
actgctgcaa acgctaattc tcttctccat ccccatgtng atattgtgta tatgtgtgag 300
ttggtnagaa tgcatcanca atctnacaat caacagcaag atgaagctag gcntgggctt 360
tcggtgaaaa tagactgtgt ctgtctgaat caaatgatct gacctatcct cggtggcaag 420
aactcttcga accgcttcct caaaggcngc tgccacattt gtggcntctn ttgcacttgt 480
ttcaaaa 487

<210> 207
<211> 332
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(332)
<223> n = A,T,C or G

<400> 207
tgaattggct aaaagactgc atttttanaa ctagcaactc ttatttcttt cctttaaaaa 60
tacatagcat taaatcccaa atcctattta aagacctgac agcttgagaa ggtcactact 120
gcatttatag gaccttctgg tggttctgct gttacntttg aantctgaca atccttgana 180
atctttgcat gcagaggagg taaaaggtat tggattttca cagaggaana acacagcgca 240
gaaatgaagg ggccaggctt actgagcttg tccactggag ggctcatggg tgggacatgg 300
aaaagaaggc agcctaggcc ctggggagcc ca 332

<210> 208
<211> 524
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(524)
<223> n = A,T,C or G

<400> 208
agggcgtggt gcggagggcg ttactgtttt gtctcagtaa caataaatac aaaaagactg 60
gttgtgttcc ggccccatcc aaccacgaag ttgatttctc ttgtgtgcag agtgactgat 120
tttaaaggac atggagcttg tcacaatgtc acaatgtcac agtgtgaagg gcacactcac 180
tcccgcgtga ttcacattta gcaaccaaca atagctcatg agtccatact tgtaaatact 240
tttggcagaa tacttnttga aacttgcaga tgataactaa gatccaagat atttcccaaa 300
gtaaatagaa gtgggtcata atattaatta cctgttcaca tcagcttcca tttacaagtc 360
atgagcccag acactgacat caaactaagc ccacttagac tcctcaccac cagtctgtcc 420
tgtcatcaga caggaggctg tcaccttgac caaattctca ccagtcaatc atctatccaa 480
aaaccattac ctgatccact tccggtaatg caccaccttg gtga 524

<210> 209
<211> 159
<212> DNA
<213> Homo sapien

<400> 209
gggtgaggaa atccagagtt gccatggaga aaattccagt gtcagcattc ttgctccttg 60
tggccctctc ctacactctg gccagagata ccacagtcaa acctggagcc aaaaaggaca 120
caaaggactc tcgacccaaa ctgccccaga ccctctcca 159

<210> 210
<211> 256
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(256)
<223> n = A,T,C or G

<400> 210
actccctggc agacaaaggc agaggagaga gctctgttag ttctgtgttg ttgaactgcc 60
actgaatttc tttccacttg gactattaca tgccanttga gggactaatg gaaaaacgta 120
tggggagatt ttanccaatt tangtntgta aatggggaga ctggggcagg cgggagagat 180
ttgcagggtg naaatgggan ggctggtttg ttanatgaac agggacatag gaggtaggca 240
ccaggatgct aaatca 256

<210> 211
<211> 264
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(264)
<223> n = A,T,C or G

<400> 211
acattgtttt tttgagataa agcattgaga gagctctcct taacgtgaca caatggaagg 60
actggaacac atacccacat ctttgttctg agggataatt ttctgataaa gtcttgctgt 120
atattcaagc acatatgtta tatattattc agttccatgt ttatagccta gttaaggaga 180
ggggagatac attcngaaag aggactgaaa gaaatactca agtnggaaaa cagaaaaaga 240
aaaaaaggag caaatgagaa gcct 264

<210> 212
<211> 328
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(328)
<223> n = A,T,C or G

<400> 212
acccaaaaat ccaatgctga atatttggct tcattattcc canattcttt gattgtcaaa 60
ggatttaatg ttgtctcagc ttgggcactt cagttaggac ctaaggatgc cagccggcag 120
gtttatatat gcagcaacaa tattcaagcg cgacaacagg ttattgaact tgcccgccag 180

ttnaatttca ttcccattga cttgggatcc ttatcatcag ccagagagat tgaaaattta 240
cccctacnac tctttactct ctgganaggg ccagtggtgg tagctataag cttggccaca 300
ttttttttc ctttattcct ttgtcaga 328

<210> 213
<211> 250
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(250)
<223> n = A,T,C or G

<400> 213
acttatgagc agagcgacat atccnagtgt agactgaata aaactgaatt ctctccagtt 60
taaagcattg ctcactgaag ggatagaagt gactgccagg agggaaagta agccaaggct 120
cattatgcca aagganatat acatttcaat tctccaaact tcttcctcat tccaagagtt 180
ttcaatattt gcatgaacct gctgataanc catgttaana aacaaatatc tctctnacct 240
tctcatcggt 250

<210> 214
<211> 444
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(444)
<223> n = A,T,C or G

<400> 214
acccagaatc caatgctgaa tatttggctt cattattccc agattctttg attgtcaaag 60
gatttaatgt tgtctcagct tgggcacttc agttaggacc taaggatgcc agccggcagg 120
tttatatatg cagcaacaat attcaagcgc gacaacaggt tattgaactt gcccgccagt 180
tgaatttcat tcccattgac ttgggatcct tatcatcagc canagagatt gaaaatttac 240
ccctacgact ctttactctc tggagagggc cagtggtggt agctataagc ttggccacat 300
ttttttttcc tttattcctt tgtcagagat gcgattcatc catatgctan aaaccaacag 360
agtgactttt acaaaattcc tataganatt gtgaataaaa ccttacctat agttgccatt 420
actttgctct ccctaatata cctc 444

<210> 215
<211> 366
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(366)
<223> n = A,T,C or G

<400> 215
acttatgagc agagcgacat atccaagtgt anactgaata aaactgaatt ctctccagtt 60
taaagcattg ctcactgaag ggatagaagt gactgccagg agggaaagta agccaaggct 120
cattatgcca aagganatat acatttcaat tctccaaact tcttcctcat tccaagagtt 180
ttcaatattt gcatgaacct gctgataagc catgttgaga aacaaatatc tctctgacct 240
tctcatcggt aagcagaggc tgtaggcaac atggaccata gcgaanaaaa aacttagtaa 300
tccaagctgt tttctacact gtaaccaggt ttccaaccaa ggtggaaatc tcctatactt 360